

UNIVERSIDADE FEDERAL DO PARANÁ

GUSTAVO BERNARDI PEREIRA

MINERAÇÃO DE PROCESSOS NA UNIDADE DO LABORATÓRIO DE ANÁLISES  
CLÍNICAS DE UM HOSPITAL ESCOLA TERCIÁRIO

CURITIBA

2019

GUSTAVO BERNARDI PEREIRA

MINERAÇÃO DE PROCESSOS NA UNIDADE DO LABORATÓRIO DE ANÁLISES  
CLÍNICAS DE UM HOSPITAL ESCOLA TERCIÁRIO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção, Setor de Tecnologia da Universidade Federal do Paraná, como requisito parcial para a obtenção do título de Mestre em Engenharia de Produção, Área de concentração em Inovação em Projetos, Produtos e Processos.

Orientador: Prof Dr Marcell Mariano Corrêa Maceno

CURITIBA

2019

Catálogo na Fonte: Sistema de Bibliotecas, UFPR  
Biblioteca de Ciência e Tecnologia

---

P436m    Pereira, Gustavo Bernardi

Mineração de processos na unidade do laboratório de análises clínicas de um hospital escola terciário [Recurso eletrônico] / Gustavo Bernardi Pereira – Curitiba, 2019.

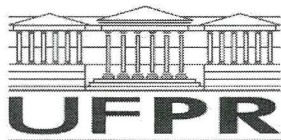
Dissertação (mestrado) - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-graduação em Engenharia de Produção.  
Orientador: Prof. Dr. Marcell Mariano Corrêa Maceno

1. Gestão de processo. 2. Administração hospitalar. 3. Engenharia de Produção. I. Universidade Federal do Paraná. II. Maceno, Marcell Mariano Corrêa. III. Título.

CDD 650.40

---

Bibliotecária: Vilma Machado CRB-9/1563



MINISTÉRIO DA EDUCAÇÃO  
SETOR SETOR DE TECNOLOGIA  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO ENGENHARIA DE  
PRODUÇÃO - 40001016070P1


## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA DE PRODUÇÃO da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **GUSTAVO BERNARDI PEREIRA** intitulada: **Mineração de Processos na Unidade do Laboratório de Análises Clínicas de um Hospital Escola Terciário**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

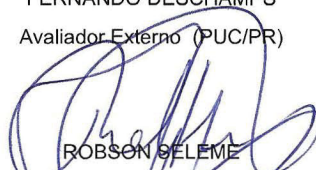
A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 20 de Fevereiro de 2019.

  
MARCELL MARIANO CORRÊA MACENO  
Presidente da Banca Examinadora (UFPR)

  
EDUARDO ALVES PORTELA SANTOS  
Avaliador Externo (PUC/PR)

  
FERNANDO DESCHAMPS  
Avaliador Externo (PUC/PR)

  
ROBSON SELEME  
Avaliador Interno (UFPR)

## **AGRADECIMENTOS**

Aos meus pais Luis e Rita pela paciência e dedicação em todas as etapas da minha vida. Da mesma forma, agradeço meus irmãos Augusto e Matheus por sempre estarem ao meu lado.

Ao meu orientador Prof. Marcell Mariano Corrêa Maceno pela humanidade, disposição e atenção, me ajudando a nortear os passos ao longo do mestrado.

À equipe do Hospital de Clínicas, que não poupou esforços para que a presente dissertação pudesse ser realizada. Agradecimento especial para: Mazé (CEP), Prof. Adonis (Direção), Gerson (Qualidade), Marisol (ULAC), Suzete (Garantia da qualidade) e Angela (Informática).

Ao Rudi Nicks da Fluxicon, que gentilmente cedeu licenças ilimitadas do Disco para a realização do projeto e auxiliou no processo de interpretação e processamento inicial dos dados.

Ao Prof. Eduardo Alves Portela Santos e Prof. Fernando Deschamps pelas incríveis aulas de MAP e BPM que ajudaram a compreender o campo e definir os rumos da pesquisa. Agradeço também a disponibilidade do Prof. Portela em transmitir seu vasto conhecimento em mineração de processos, sem o qual, não teria sido possível desenvolver a presente dissertação.

Ao Prof. Robson Seleme pelo apoio como coordenação do curso ao longo do mestrado.

À secretária do PPGEF Rafaella Laureano Scaramella Figueira da Silva pela bondade nas suas ações, que tornaram o mestrado um período mais agradável.

Aos colegas da turma de 2017 do PPGEF que, apesar de não conseguirem marcar um churrasco, fizeram com que boas risadas marcassem esses dois anos de parceria e muitos mais que estão por vir: André, Andreia, Anis, Carlos Brasil, Carlos Flesch, Daniel Ayub, Daniel Oliveira, Danielle, Gregori, Henrique, Iris, Jefferson, João Vitor, Lívia, Majed, Matheus, Natalia, Nicolas, Paulo, Ricardo, Rodrigo, Tarley, Thiago, Tiago e Vitor.

A todos que, direta ou indiretamente, contribuíram para a realização desse trabalho.

*Não sofreremos com os acontecimentos  
de nossas vidas, mas com nosso  
julgamento sobre eles*

*Epictetus*

## RESUMO

O mapeamento dos processos no ambiente de saúde proporciona diversos benefícios gerenciais, que são refletidos na qualidade do atendimento dos pacientes. Dentre as maneiras de fazê-lo, um método denominado “mineração de processos” vem sendo usado em diversos contextos e tem apresentado resultados interessantes. Porém, existe uma carência de estudos voltados à padronização do processo de aplicação da mineração de processos. Tentando cobrir tal lacuna, o presente estudo elaborou uma metodologia de aplicação da mineração de processos na área da saúde, intitulada *Process Mining Project Methodology in Healthcare* (PM<sup>2</sup>HC). Essa metodologia foi desenvolvida ao longo de uma série de etapas envolvendo levantamentos bibliográficos sobre as metodologias de aplicação da mineração de processos no âmbito geral e sobre aplicações da mineração de processos em estudos de caso da área da saúde. Foram utilizadas as bases Scopus, Web of Science, Science Direct, PubMed e Google Scholar. A partir dos artigos que atenderam os critérios de qualidade estabelecidos (relativos ao local de publicação, número de citações, dentre outros), foram identificados elementos que contribuíram para o processo de aplicação. Com base nesses elementos, as metodologias de aplicação da mineração de processos foram avaliadas e foi então selecionado o PM<sup>2</sup> como modelo base. Além dos elementos extraídos dos artigos, alguns aspectos mais técnicos também foram adicionados para deixar o PM<sup>2</sup> adaptado para o ambiente da saúde, tornando-se PM<sup>2</sup>HC. Por fim, o PM<sup>2</sup>HC foi aplicado na Unidade do Laboratório de Análises Clínicas do Complexo Hospital de Clínicas da Universidade Federal do Paraná para poder ser avaliado em uma situação com modelos de processo e dados reais. Os resultados mostram que o PM<sup>2</sup>HC proporciona uma grande proximidade dos tomadores de decisão (*Stakeholders*) com os analistas de mineração, permitindo que os objetivos sejam melhor definidos e, ao final do projeto, seja possível confrontar o resultado obtido com o esperado (com o uso de KPIs, por exemplo). Além disso, por tratar de aspectos técnicos ao longo das etapas do projeto, o PM<sup>2</sup>HC traz maior segurança para analistas de mineração de processos que não possuam tanta experiência na área.

**Palavras-chave:** Mineração de processos, Hospital escola terciário, Disco, Área da saúde.

## ABSTRACT

Process mapping in the healthcare environment provides several managerial benefits, which are reflected in the quality of patient care. Among the ways to map the processes, a method called "process mining" has been used in several contexts and has presented interesting results. However, there is a lack of studies focused on the standardisation of the process mining application process. In order to bridge this gap, the present study developed a methodology for the application of process mining in healthcare entitled Process Mining Project Methodology in Healthcare (PM<sup>2</sup>HC). This methodology was developed over a series of steps involving bibliographical reviews on the methodologies of application of process mining in the general and on applications of process mining in health case studies. The used databases were Scopus, Web of Science, Science Direct, PubMed and Google Scholar. From the articles that met the established quality criteria (relative to the publication source, number of citations, among others), elements that contributed to the deployment process application were identified. Based on these elements, the process mining application methodologies were evaluated and the PM<sup>2</sup> was selected as the base model. In addition to the elements extracted from the articles, some more technical aspects were also added to make the PM<sup>2</sup> adapted to the health environment, becoming PM<sup>2</sup>HC. Finally, the PM<sup>2</sup>HC was applied in the Clinical Analysis Laboratory Unit of the Tertiary Clinical Hospital Complex of the Federal University of Paraná to be evaluated in a situation with real process models and data. The results show that PM<sup>2</sup>HC provides more proximity between decision makers (Stakeholders) and mining analysts, allowing the objectives to be better defined and, at the end of the project, it is possible to compare the result obtained with the expected of KPIs, for example). In addition, by addressing technical aspects throughout the project stages, the PM<sup>2</sup>HC provides guidelines for process mining analysts who do not have as much experience in the field.

**Keywords:** process mining, tertiary school hospital, Disco, healthcare.



## LISTA DE SIGLAS

BP	–	<i>Business Process</i>
BPM	–	<i>Business Process Management</i>
BPMS	–	<i>Business Process Management Systems</i>
BPR	–	<i>Business Process Re-engineering</i>
CAAE	–	Certificado de Apresentação para Apreciação Ética
CEP	–	Comitê de Ética em Pesquisa
EBSERH	–	Empresa Brasileira de Serviços Hospitalares
FSM	–	<i>Finite-state machines</i>
HU	–	Hospitais universitário
ID	–	<i>Identifier/Identificador</i>
KPI	–	<i>Key Performance Indicator</i>
PDCA	–	<i>Plan, Do, Check, Action</i>
PDM	–	<i>Product Data Model</i>
PM <sup>2</sup>	–	<i>Process Mining Project Methodology</i>
PM <sup>2</sup> HC	–	<i>Process Mining Project Methodology in Healthcare</i>
RAM	–	<i>Random Access Memory</i>
RS	–	Regional de Saúde
RSL	–	Revisão Sistemática da Literatura
HIS	–	Sistema de Informação Hospitalar
SUS	–	Sistema Único de Saúde
SQL	–	<i>Structured Query Language</i>
TI	–	Tecnologia da Informação
UPME	–	Unidade de Processamento de Materiais Esterilizados
ULAC-CHC	–	Unidade do Laboratório de Análises Clínicas do Complexo Hospital de Clínicas
UFPR	–	Universidade Federal do Paraná
V&V	–	Verificar e validar
WfMS	–	<i>Workflow Management Systems</i>

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>11</b>
1.1 JUSTIFICATIVA .....	13
1.2 OBJETIVOS .....	14
1.3 DELIMITAÇÃO DO TRABALHO .....	15
1.4 ESTRUTURA DO TRABALHO .....	15
<b>2. CONCEITOS FUNDAMENTAIS .....</b>	<b>17</b>
2.1 <i>BUSINESS PROCESS</i> .....	17
2.1.1 <i>Business Process Management</i> .....	18
2.1.2 <i>Business Process Management Systems</i> .....	21
2.2 <i>WORKFLOW MANAGEMENT SYSTEMS</i> .....	22
2.3 MINERAÇÃO DE PROCESSOS .....	24
2.3.1 Ferramentas de mineração de processos .....	25
2.3.2 Mineração de processos na área da saúde.....	26
<b>3. METODOLOGIA DA PESQUISA.....</b>	<b>28</b>
3.1 CARACTERIZAÇÃO DA PESQUISA .....	28
3.2 PROTOCOLO DE PESQUISA .....	29
3.2.1 Pré-desenvolvimento.....	30
3.2.2 Desenvolvimento .....	32
3.2.3 Aplicação e avaliação.....	32
3.2.4 Conclusão .....	32
<b>4 PRÉ-DESENVOLVIMENTO .....</b>	<b>33</b>
4.1 IDENTIFICAÇÃO DO PROBLEMA .....	33
4.2 REVISÃO SISTEMÁTICA DA LITERATURA .....	33
4.2.1 Estado da arte e aplicações da mineração de processos .....	34
4.2.1.1 Identificação dos artigos .....	34
4.2.1.2 Leitura diagonal para identificar se é elegível .....	36

4.2.1.3 Leitura dos artigos completos.....	36
4.2.1.4 Análise transversal .....	38
4.2.2 Identificação de modelos de mineração de processos.....	39
4.2.2.1 Identificação dos artigos.....	40
4.2.2.2 Leitura diagonal para identificar se é elegível .....	41
4.2.2.3 Leitura dos artigos completos.....	41
4.2.2.4 Análise transversal .....	42
<b>5 DESENVOLVIMENTO .....</b>	<b>55</b>
5.1 SELEÇÃO DA METODOLOGIA.....	55
5.1.1 <i>Process Mining Process Methodology</i> (PM <sup>2</sup> ).....	58
5.2 APRESENTAÇÃO DA METODOLOGIA PM <sup>2</sup> HC .....	68
<b>6 APLICAÇÃO .....</b>	<b>72</b>
6.1 O COMPLEXO HOSPITAL DE CLÍNICAS DA UNIVERSIDADE FEDERAL DO PARANÁ.....	72
6.2 APLICAÇÃO.....	73
6.2.1 Planejamento .....	74
6.2.1.1 Seleção de processos .....	74
6.2.1.2 Identificação das questões de pesquisa.....	75
6.2.1.3 Composição da equipe de projeto .....	76
6.2.2 Extração .....	76
6.2.2.1 Determinação do escopo.....	76
6.2.2.2 Extração de dados de eventos .....	77
6.2.2.3 Transferência de conhecimento do processo.....	78
6.2.3 Processamento de dados.....	78
6.2.3.1 Criação de visualizações.....	79
6.2.3.2 Agregação de eventos.....	81
6.2.3.3 Enriquecimento dos <i>logs</i> .....	81

6.2.3.4 Filtragem de <i>logs</i> .....	81
6.2.4 Mineração e Análise .....	83
6.2.4.1 Descoberta de Processos .....	83
6.2.4.2 Verificação de conformidade .....	88
6.2.4.3 Análise do processo .....	92
6.2.4.4 Aprimoramento .....	93
6.2.5 Avaliação .....	93
6.2.5.1 Diagnosticar .....	93
6.2.5.2 Verificar e validar (V&V) .....	94
6.2.6 Melhoria de Processos e Suporte .....	94
6.2.6.1 Implementar melhorias .....	94
6.2.6.2 Suporte de operações .....	94
6.3 AVALIAÇÃO .....	95
6.4 RESULTADO FINAL DO PM <sup>2</sup> HC .....	97
6.5 APRENDIZADO COM A APLICAÇÃO DO PM <sup>2</sup> HC .....	100
6.5.1 Compilação dos achados .....	100
<b>7 CONSIDERAÇÕES FINAIS .....</b>	<b>102</b>
<b>REFERÊNCIAS .....</b>	<b>104</b>
<b>ANEXO 1 - ROLE ACTIVITY DIAGRAM (RAD) .....</b>	<b>115</b>

## 1 INTRODUÇÃO

As organizações de saúde buscam prestar serviços aos seus pacientes prezando pela qualidade, tempo e funcionalidade. Mas isso não é tão simples quanto parece (GUPTA, 2007).

O tema “saúde” é constantemente discutido na esfera pública, em especial por afetar a qualidade de vida das pessoas e a gestão do Governo. Dessa forma, a alocação de recursos nessa área é frequente, buscando promover a melhoria das condições de vida da população (INSPER, 2012).

No Brasil, o Sistema Único de Saúde (SUS) oferece atendimento universal e gratuito de saúde, porém, apresenta diversos problemas estruturais que conduzem à precariedade dos serviços prestados (INSPER, 2012). De forma a solucionar tais problemas, a gestão de recursos do SUS tem tido a inserção gradativa de tecnologia, buscando informatizar seus serviços. Tais ações têm o objetivo de encontrar pontos onde estão desperdícios de recursos. A visão é a criação de uma política de resultados no ministério (BRASIL. MINISTÉRIO DA SAÚDE, 2016).

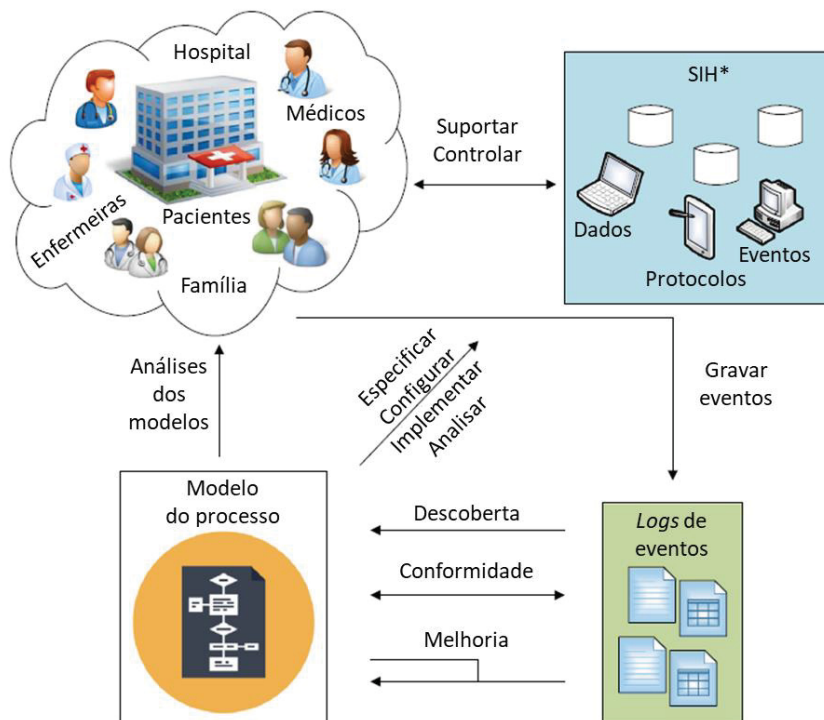
Seguindo essa tendência tecnológica, é interessante ressaltar que os centros de atenção à saúde, de modo geral, possuem sistemas que fazem o registro das atividades executadas (o termo *log* é a denominação adotada usualmente para os registros de eventos e, portanto, será utilizado no presente trabalho). Tais *logs* são úteis no delineamento e análise dos processos desenvolvidos. Tais informações podem ser utilizadas para automaticamente fornecer um modelo do processo que pode ser utilizado para a descoberta de suas forças e fraquezas. Essa abordagem recebe o nome de mineração de processos (GUPTA, 2007).

A mineração de processos é uma área de estudo que se situa entre inteligência computacional e mineração de dados, de um lado, e modelagem e análise de processos, de outro lado (VAN DER AALST et al., 2012). Seu principal objetivo é extrair conhecimento de dados que foram gerados e armazenados em sistemas de dados corporativos para criar *logs* de eventos. Os *logs* de eventos são conjuntos de registros em que todas as atividades executadas para uma instância de processo particular são registradas (ROJAS et al., 2016). O conhecimento adquirido dos *logs* é usado para descobrir, monitorar e melhorar processos reais (SONG; VAN DER AALST, 2008).

Sua aplicação em saúde encontra dificuldades devido à demanda intrínseca que os processos de saúde têm em relação à informação e ao conhecimento (LENZ; REICHERT, 2007). Porém, está se transformando em uma realidade desde a década de 2000, uma vez que as instituições de saúde adotaram sistemas de gerenciamento capazes de gerar *logs* das diferentes atividades realizadas (RIZ; SANTOS; LOURES, 2016).

A FIGURA 1 mostra um esboço geral da aplicação da mineração de processos em cuidados de saúde de acordo com Rojas (2016). As atividades realizadas por recursos hospitalares (por exemplo, médico, enfermeiro, técnico, etc.) com o objetivo de cuidar de um paciente são armazenadas em um Sistema de Informação Hospitalar (SIH), que é composto de bancos de dados, sistemas, protocolos e eventos. Essas atividades são gravadas como *logs* de eventos para suporte, controle e análise posterior. Os modelos de processos são criados para analisar e sugerir a ordem em que os profissionais de saúde devem realizar suas atividades dentro de um determinado processo ou avaliar criticamente o *design* do processo. Da mesma forma, os modelos de processo também são usados para apoiar o desenvolvimento do SIH, por exemplo, para entender como o sistema de informações deve suportar a execução do processo (ROJAS et al., 2016).

FIGURA 1 - MINERAÇÃO DE PROCESSOS EM SAÚDE



FONTE: (ROJAS et al., 2016, tradução nossa)  
\*Sistemas de Informação Hospitalar

O uso desse conceito permite que seja conhecido, por exemplo, o fluxo de trabalho do tratamento, que é uma informação importante para os profissionais da área da saúde. Este fluxo descreve todos os processos de interação entre o ambiente e os agentes (pacientes, médicos, enfermeiros, etc.) e tem recebido crescente atenção na área da saúde para a automação dos processos (WANG, 2012). Além disso, tanto o governo quanto as companhias de seguro têm aumentado a pressão sobre os hospitais para que o trabalho seja feito de modo mais eficiente possível. Essa pressão incorpora também, a tendência de crescimento da demanda por serviços na área de *healthcare* (MANS et al., 2009). Dessa forma, as ferramentas de mineração de processos podem ser utilizadas para o processo de diagnóstico e tratamento não apenas de pacientes (ex: análise das etapas pelas quais um determinado caso passou), mas do sistema de saúde como um todo (ex: exploração dos caminhos clínicos de todos os pacientes que deram entrada em uma seção).

Tal multiplicidade de aplicações faz com que a maneira de como a mineração de processos na saúde é aplicada não possua diretrizes comuns muito bem definidas, apesar de existirem metodologias para a mineração de processos no âmbito geral (VAN DER HEIJDEN, 2011; VAN ECK et al., 2015).

## 1.1 JUSTIFICATIVA

A contribuição que a presente pesquisa apresenta encontra-se na adaptação de uma metodologia de aplicação da mineração de processos específica para a área da saúde, com validação por meio de estudo de caso em um hospital escola<sup>1</sup> terciário<sup>2</sup>.

Tal necessidade foi suscitada pela falta de trabalhos voltados à padronização do processo de aplicação da mineração de processos (a maioria dos trabalhos busca

---

<sup>1</sup> Os hospitais escola são centros de formação de recursos humanos e de desenvolvimento de tecnologia para a área de saúde (BRASIL, 2012). Isso significa que desempenham pesquisa científica enquanto realizam os tratamento e capacitação de profissionais de saúde, sendo mantidos e/ou atuando em colaboração com as mesmas (MELDAU, 2013).

<sup>2</sup> No Brasil, existem quatro níveis de atenção à saúde: no nível **primário**, as ações compreendem promoção, prevenção e proteção à saúde, tendo no Programa Saúde da Família sua principal estratégia de organização da atenção básica (assim como as Unidades Básicas de Saúde). No nível **secundário**, a prática é pautada pelo atendimento e tratamento por especialidades (Unidades de Pronto Atendimento). O nível **terciário** contempla as internações/hospitalizações (hospitais de grande porte) e o **quaternário** envolve as práticas de reabilitação (DAMOUS, I.; ERLICH, H., 2017).

a padronização do fluxo de pacientes) (FEI; MESKENS; MOREAU, 2009; FUNKNER; YAKOVLEV; KOVALCHUK, 2017; STAAL, 2010). Além disso, as poucas metodologias estabelecidas que se dispõem a tratar do assunto (ex: L\*, PDM, dentre outros), são voltadas aos projetos de mineração de processos no âmbito generalista e não aplicáveis especificamente à área da saúde (BOZKAYA; GABRIELS; WERF, 2009).

Neste contexto, a carência de metodologias de mineração de processos específicas da área da saúde faz com que a curva de aprendizado para que se possa desenvolver um projeto de mineração de processos torne-se muito complexa e trabalhosa. Por esses motivos, o processo de tentativa e erro encontra-se tão presente nessa área; revisando e realizando ajustes até que os resultados obtidos sejam coerentes e significativos (GANDULFO, 2016). Além disso, poucas das metodologias para aplicação geral da mineração de processos encorajam explicitamente a análise iterativa e/ou a associação com o conhecimento tácito contido nas pessoas envolvidas no processo a ser estudado (SURIADI et al., 2013; VAN ECK et al., 2015).

Dessa forma, uma metodologia específica de aplicação da mineração de processos em saúde permite que o desenvolvimento dos projetos seja mais uniforme e consistente. Isso permite que futuras análises utilizando ferramentas de mineração de processos possam ser estruturadas de forma a atender melhor às demandas individuais de cada caso. Além de fornecer diretrizes práticas para auxiliar analistas inexperientes a superarem obstáculos encontrados.

Além disso, em última instância, o presente projeto apresenta benefícios para a população geral por permitir a aplicação de ferramentas na área da saúde. Isso permite que os processos de gestão sejam diagnosticados e revistos para atender às demandas do setor.

## 1.2 OBJETIVOS

### – Objetivo Geral

O objetivo geral da presente dissertação foi elaborar uma metodologia de aplicação da mineração de processos na área da saúde.



## – Objetivos Específicos

A fim de atingir o objetivo geral, os objetivos específicos são:

- Apresentar o estado da arte da mineração de processos em ambientes de saúde;
- Selecionar metodologias de mineração de processos no âmbito geral que sejam factíveis e adequadas para serem aplicadas na saúde;
- Estabelecer elementos importantes em um projeto de mineração de processos
- Avaliar a metodologia por meio de um estudo de caso.

### 1.3 DELIMITAÇÃO DO TRABALHO

Essa pesquisa delimita-se a adaptar uma metodologia de aplicação da mineração de processos na área da saúde, tendo como base o estudo do processamento dos exames requisitados na Unidade do Laboratório de Análises Clínicas (ULAC) do Complexo Hospital de Clínicas da Universidade Federal do Paraná (CHC-UFPR), localizada em Curitiba – Paraná. A janela de tempo considerada foi de 1º de janeiro de 2016 a 31 de dezembro de 2017. O CHC-UFPR é o terceiro maior hospital-escola do Brasil e é referência nacional em tratamentos de média e alta complexidade.

### 1.4 ESTRUTURA DO TRABALHO

Esta pesquisa está organizada em sete capítulos. No primeiro e presente capítulo, **Introdução**, são descritos a introdução ao tema, a justificativa, o objetivo geral e os objetivos específicos, a delimitação da pesquisa e, por fim, a estrutura do trabalho.

No capítulo 2 – **Conceitos Fundamentais**, são apresentados os principais conceitos utilizados nesse projeto, percorrendo temas como: *Business Process*, *Workflow Management* e Mineração de Processos.

O capítulo 3 – **Metodologia da pesquisa**, destina-se, primeiramente, a descrever a caracterização da pesquisa, quanto a natureza, o objetivo e o

procedimento de pesquisa. Em seguida, delinea o protocolo de pesquisa com as etapas e passos percorridos para alcançar os objetivos traçados.

No capítulo 4 – **Pré-Desenvolvimento**, são apresentadas as duas revisões da literatura que foram desenvolvidas para que fossem apresentados o estado da arte da mineração de processos em um âmbito geral e suas aplicações em saúde.

No capítulo 5 – **Desenvolvimento**, a metodologia base é selecionada com base em um conjunto de critérios definidos como desejáveis. Em seguida, ela é complementada com os comentários e observações presentes nas aplicações elencadas no capítulo anterior.

No capítulo 6 – **Aplicação**, é apresentada a instituição em que a aplicação foi feita, é realizada a aplicação da metodologia e é feita a sua avaliação.

Por fim, no capítulo 7 – **Considerações finais**, são tecidas as conclusões acerca do trabalho e sugeridas pesquisas futuras.

## 2. CONCEITOS FUNDAMENTAIS

Este capítulo tem como objetivo fornecer conceitos para o melhor entendimento da pesquisa, introduzindo os conceitos-chave (e seus desmembramentos específicos).

Estão apresentados: *Business Process* (2.1), *Business Process Management* (2.1.1) *Business Process Management Systems* (2.1.2), *Workflow Management Systems* (2.2), Mineração de Processos (2.3), Ferramentas de Mineração de Processos (2.3.1) e Mineração de Processos na área da Saúde (2.3.2).

### 2.1 BUSINESS PROCESS

O modelo de negócios dos anos 2010-2020 apresenta um contexto de competição desenvolvido entre empresas dos diversos ramos que vem associado ao crescimento contínuo das expectativas dos clientes. Tal associação pressiona as companhias a melhorarem seus níveis de eficiência ao mesmo tempo em que mantém flexibilidade suficiente para adequarem-se às mudanças aceleradas no ambiente. Surge então a necessidade de agilizar suas operações ao mesmo tempo em que os *Business Process* (BP)<sup>3</sup> são integrados ao longo dos silos funcionais (SIDOROVA; ISIK, 2010).

A necessidade de integração transversal de funções obriga as empresas a adotarem uma abordagem orientada a processos para que suas operações possam ser gerenciadas. Fundamentalmente, um processo é uma função organizacional que “encapsula a interdependência de tarefas, papéis, pessoas, departamentos, funções etc., que é necessária para fornecer um produto ou serviço a um cliente (interno ou externo)” (EARL, 1994, p. 13 tradução nossa). Um processo tem início, fim e elementos de entrada e saída claramente identificados: uma estrutura de ação (DAVENPORT, 1993).

Conforme histórico apresentado por Sidorova (2010), a ideia de *Business Process* pode ser remontada partindo do exemplo da fábrica de alfinetes que Adam Smith apresentou em 1776 e de sua descrição da divisão do trabalho (SMITH, 1976). Posteriormente, Taylor (1911) desenvolveu as bases da administração científica,

---

<sup>3</sup> Processos de negócio

evoluindo os conceitos de BP. A partir dos anos 50, houve a expansão da abordagem estatística para os *Business Processes* com os trabalhos de Shewhart, Juran e Deming na gestão da qualidade (BEIMBORN; JOACHIM, 2011; MÜNSTERMANN; ECKHARDT; WEITZEL, 2010). Porém, até este momento, os autores mencionados mantiveram seu foco quase que exclusivamente nos processos de produção e manufatura.

Apenas nos anos 1980 foi quando a visão dos processos se tornou amplamente adotada, com uso do *Business Process Re-engineering* (BPR)<sup>4</sup>; provendo aos indivíduos envolvidos uma visão holística das atividades que são realizadas dentro das unidades de negócios e, assim, melhorando a efetividade organizacional (CHILDE et al., 1995). Tais aspectos foram expandidos com a quarta (1980-1995) e a quinta (1995-atual) gerações do gerenciamento de dados, com a presença de bases de dados relacionais, processamento de bases de dados em paralelo, programação orientada ao objeto sendo utilizados por empresas. Tais aprimoramentos, associados à evolução acelerada do poder de processamento da computação e comunicação, tornaram possível a digitalização total dos processos de negócio organizacionais (GRAY, 2007). O crescente interesse no aprimoramento dos BPs associado à evolução tecnológica acelerada fez com que as mudanças dos BPs fossem orientadas pela tecnologia, buscando a integração dos processos por meio de sistemas de planejamento e, mais recentemente, com o uso de sistemas de *Business Process Management* (BPM)<sup>5</sup> (CAVERLEE et al., 2007).

### 2.1.1 *Business Process Management*

O *Business Process Management* é um modelo que busca capturar e apoiar a infraestrutura de tecnologia de informações empresariais (CAVERLEE et al., 2007). Tal definição pode ser expandida com o que é apresentado por Trkman (2010, p. 125 tradução nossa), que coloca que o BPM é um esforço realizado pelas organizações para “analisar e melhorar continuamente as atividades fundamentais, como fabricação, marketing, comunicações e outros elementos importantes das operações

---

<sup>4</sup> Reengenharia de processos de negócios

<sup>5</sup> Gerenciamento de processos de negócios

da empresa”. Ambas as visões podem ser consolidadas com o que é apresentado pela *Association of Business Process Management Professionals* (ABPMP)<sup>6</sup>:

*(...) tem o propósito de medir, monitorar, controlar atividades e administrar o presente e o futuro do negócio. Processos de gerenciamento, assim como os processos de suporte, não agregam valor diretamente para os clientes, mas são necessários para assegurar que a organização opere de acordo com seus objetivos e metas de desempenho(...)* (ABPMP, 2013, p. 49)

A abrangência e tendência dessa última definição vão ao encontro da pesquisa desenvolvida por Seethamraju e Marjanovic (2009), onde apresentam que as companhias têm modificado seu perfil de gestão de modo a adotar um modelo mais focado nos processos, ao invés da tradicional abordagem funcional. Essa mudança de paradigma, no que diz respeito à adoção do BPM, traz potenciais benefícios conforme cita Caverlee et al. (2007):

- fornece uma visão clara da estrutura do negócio, facilitando a identificação de gargalos no processo e a reutilização de infraestrutura existente;
- aumenta a qualidade dos outputs por meio de um gerenciamento uniforme e consistente do processamento de informações em subsistemas cada vez mais heterogêneos;
- automatiza a carga de trabalho de gerenciamento de baixo nível, como monitoramento de WIP (*work in progress*) ou sequenciamento de produção, permitindo que as empresas possam concentrar-se em oportunidades e atividades com maior valor agregado.

É necessária, porém, uma observação com relação aos benefícios apresentados. O BPM normalmente é introduzido em uma organização através de projetos de curto prazo que buscam resolver ineficiências específicas. Porém, é oportuno ir além da busca de vitórias momentâneas. O princípio da continuidade enfatiza que o BPM deve ser uma prática permanente que facilite ganhos contínuos de eficiência e eficácia. Estabelecer uma abordagem BPM de longo prazo e a instalação de uma mentalidade de processo de forma sustentável é importante para poder aproveitar o potencial e o valor do BPM (VOM BROCKE et al., 2014).

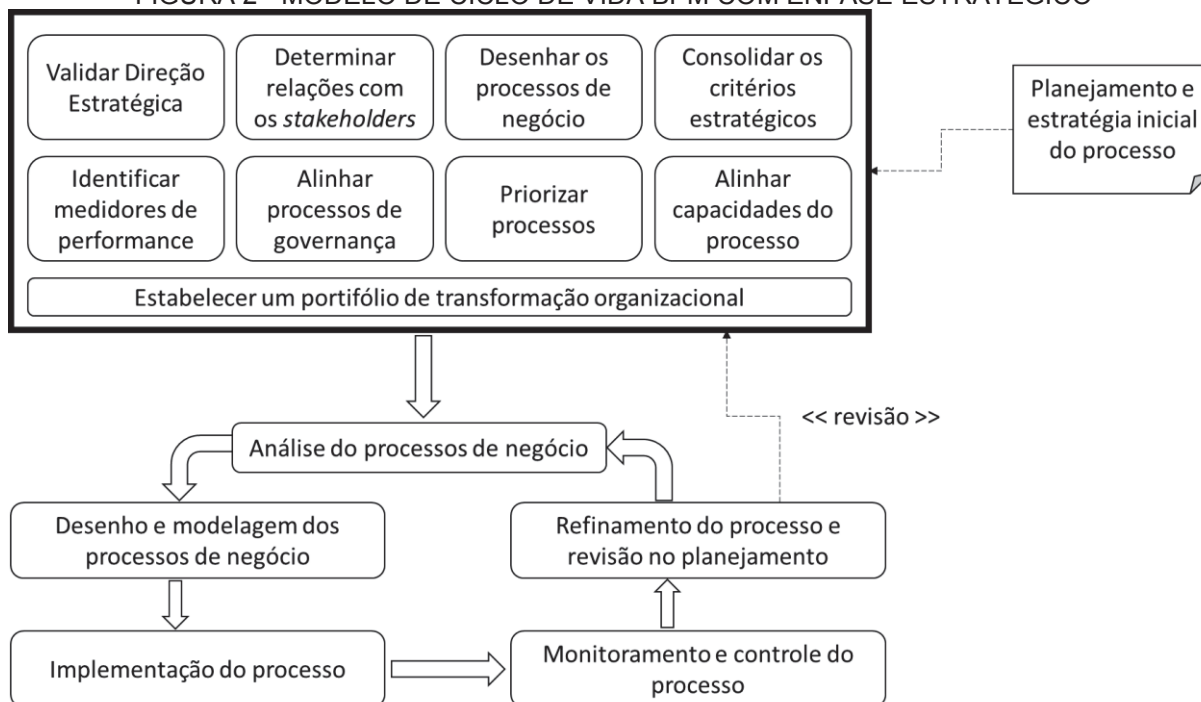
Mantendo a ênfase em um modelo estratégico de BPM, Moraes et al. (2014) realizaram uma comparação entre os sete modelos de ciclo de vida BPM mais presentes na literatura e apresentaram um modelo que busca integrá-los de modo a

---

<sup>6</sup> Associação de Profissionais de Gerenciamento de Processos de Negócios

(1) manter o alinhamento entre estratégia e processos de negócio de maneira explícita e (2) possuir boa aplicabilidade tanto no campo acadêmico quanto no prático. O modelo apresentado (FIGURA 2) é o que será adotado para fins da presente dissertação.

FIGURA 2 - MODELO DE CICLO DE VIDA BPM COM ÊNFASE ESTRATÉGICO



FONTE: Adaptado de Moraes et al. (2014)

A leitura da FIGURA 2 inicia-se com a concepção da arquitetura do processo (atividades dentro do retângulo em negrito). Nela são desenvolvidas as nove etapas iniciais do planejamento e estratégia do processo. Em seguida, o modelo segue para um ciclo que contempla as etapas de (1) análise, (2) projeto e modelagem, (3) implementação, (4) monitoramento e controle de processos e (5) revisão e refinamento. O último passo precede cada nova iteração do ciclo, que analisa as atividades associadas à estratégia para o portfólio de processos de negócios.

Para que as tarefas apresentadas possam ser desenvolvidas, é conveniente que seja feito o uso de um *Business Process Management System* (BPMS)<sup>7</sup>; enquanto um processo de negócio é um sistema sócio técnico, executado por humanos e máquinas, o BPMS é puramente o sistema técnico (DUNCAN et al., 2007).

<sup>7</sup> Sistema de gerenciamento de processos de negócio

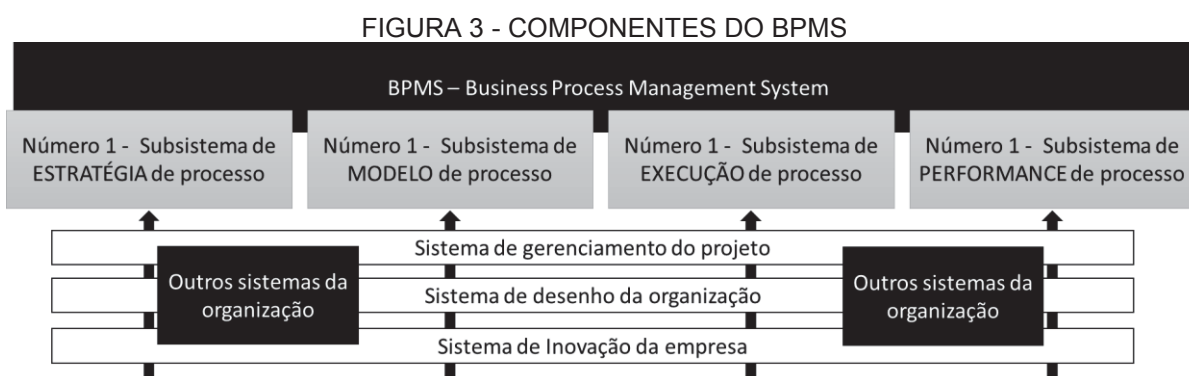
### 2.1.2 Business Process Management Systems

O BPMS pode ser definido como “um sistema de software genérico que é conduzido pelo *design* de processos explícitos para implementar e gerenciar processos de negócios operacionais.” (VAN DER AALST; HOFSTEDE; WESKE, 2003, p. 1 tradução nossa). De modo mais prático, pode ser também visto como a arquitetura organizacional que integra todas as abordagens, métodos, técnicas e aplicações tecnológicas focadas no alinhamento sistemático de estratégias e operações, bem como na construção de uma organização voltada a processos e à geração de valor (MARGHERITA, 2014).

Partindo dessas definições, é possível apresentar os quatro subsistemas contidos em um BPMS:

- estratégia de processo;
- modelo de processo;
- execução do processo;
- desempenho do processo.

Os subsistemas mencionados não compõem fases de um ciclo de vida de um processo genérico, mas sim elementos complementares de uma arquitetura orgânica que busca centralizar os processos de negócio bem como seus mecanismos no contexto da organização. Além disso, o BPMS encontra-se relacionado a outros três “sistemas” da organização (sistema de gerenciamento de projetos, o sistema de *design* da organização e o sistema de inovação empresarial) (MARGHERITA, 2014). A FIGURA 3 ilustra o BPMS e suas relações com os subsistemas e os “sistemas” da organização.



FONTE: Adaptado de Margherita (2014)

Exemplos de BPMS presentes na literatura são os *Enterprise Resources Planning* (ERPs) (VAN DER AALST; HOFSTEDE; WESKE, 2003) e os *Workflow Management Systems* (WFMS) (DUNCAN et al., 2007; GUPTA, 2007; SAYAL et al., 2002; VAN DER AALST; HOFSTEDE; WESKE, 2003). Para a presente dissertação o BPMS utilizado será o *Workflow management System* pelo fato de estar mais alinhado com o tipo de trabalho desenvolvido.

## 2.2 WORKFLOW MANAGEMENT SYSTEMS

O *Workflow*<sup>8</sup> diz respeito a tarefas, recursos e *triggers*<sup>9</sup> associados a um processo específico (GUPTA, 2007). Expandindo-se a definição, pode ser utilizada a versão adotada pela *Workflow Management Coalition*<sup>10</sup> (1996) que define workflow como a automação de um *Business Process* como todo ou em parte, durante a qual documentos, informações ou tarefas são passados de um participante para outro para que uma ação seja tomada, conforme um grupo pré-determinado de regras processuais.

Por sua vez, os *Workflow Management Systems* (WfMS)<sup>11</sup> são um conjunto de ferramentas que possibilitam o *design*, a análise e a simulação dos *Business Process* orientados ao modelo, podendo ser projetados do zero ou partindo de algum modelo pré-existente. WfMSs também fornecem recursos para monitorar a execução dos *Business Processes* e para reagir automaticamente a situações excepcionais (SAYAL et al., 2002). Esses sistemas requerem um modelo de processo e sua principal função é assegurar que todas as atividades sejam desenvolvidas na sequência correta e pelo recurso correto (SARAVANAN; RAMA SREE, 2011).

Mantendo-se a ideia de sequência de execução pelo recurso correto, Wendler & Loef (2001) acrescentam uma lista do que é possível ser feito com os WfMSs:

- modelar os processos em termos de atividades e transições de estado;
- modelar a organização em termo de unidades organizacionais e participantes no workflow;
- combinar os participantes do *workflow* e as atividades;

---

<sup>8</sup> Fluxo de trabalho

<sup>9</sup> Triggers podem ser vistos como elementos que desencadeiam outros processos/tarefas (gatilhos)

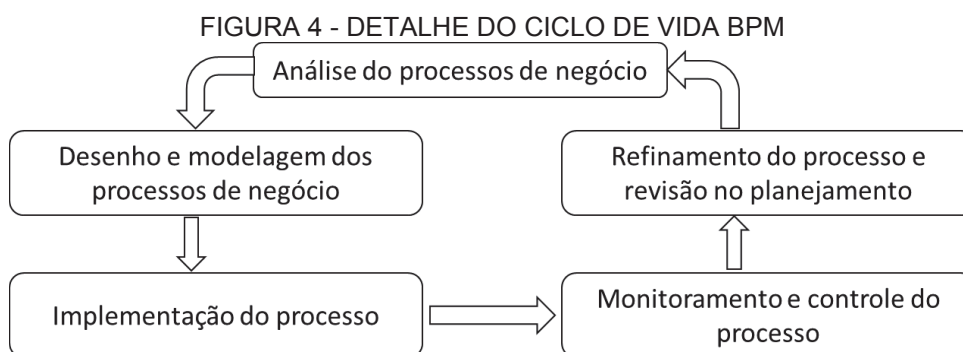
<sup>10</sup> Fundada em 1993, a *Workflow Management Coalition* (WfMC) é uma organização global de adeptos, desenvolvedores, consultores, analistas, bem como universidades e grupos de pesquisa envolvidos em *workflow* e BPM

<sup>11</sup> Sistemas de gerenciamento do fluxo de trabalho



- colocar processos em ação e fornecer listas de trabalho para os sistemas de aplicativos.

Tais itens podem ser alocados dentro do ciclo de vida BPM apresentado na seção 2.1.1 *Business Process Management*, cujo detalhe encontra-se apresentado na FIGURA 4. O foco dos WfMSs encontra-se tipicamente nas fases de *design* e implementação.



FONTE: Adaptado de Moraes et al. (2014)

A fase de análise agrega atividades destinadas a alinhar os objetivos de negócios com seus processos, seja para estabelecê-los ou atualizá-los, e com base no escopo, são aplicadas técnicas para mapear o contexto comercial através de entrevistas, análises documentais, simulações ou outros instrumentos de prospecção (MORAIS et al., 2014).

Na fase de *design* do processo, os processos operacionais são projetados ou redesenhados conforme as regras definidas para compartilhar informações entre os grupos funcionais, o *design* físico e a infraestrutura de TI. A fase de implementação do sistema deve ser vista como uma atividade de “orquestração”, onde o projeto feito na fase anterior é implementado configurando um sistema de informação como o *Workflow Management System* (WFMS). O monitoramento e controle de processos trata de ajustes de recursos para garantir objetivos de processo através de medições de desempenho e avaliação. Por fim, a etapa de refinamento está associada a mudanças organizacionais, melhoria contínua e atividades de otimização em busca da eficiência e eficácia dos processos implementados na organização (MORAIS et al., 2014).

Quando é realizada a mineração do *workflow*, também conhecido como Mineração de Processos, é obtido um modelo que reflete o fluxo de trabalho existente, sendo valioso para a compreensão e redesenho dos processos (BLUM et al., 2008).

## 2.3 MINERAÇÃO DE PROCESSOS

A mineração de processos é uma área de estudo recente (VAN DER AALST; HOFSTEDE; WESKE, 2003; VAN DER AALST; WEIJTERS, 2004) e situa-se entre duas áreas (1) inteligência computacional e mineração de dados e (2) modelagem e análise de processos (VAN DER AALST et al., 2012). Seu objetivo primário é a extração de conhecimento a partir dos dados que são gerados e armazenados em sistemas de dados de empresas na forma de *logs* de eventos. Os *logs* de eventos são conjuntos de registros nos quais todas as atividades de um processo são registradas (ROJAS et al., 2016). Apesar de existir muita semelhança entre a mineração de dados com a mineração de processos a principal diferença é que: enquanto a primeira busca explicitar tendências nos dados e a relação entre seus atributos, a segunda tem como foco a compreensão de como determinado processo é executado, baseando-se na análise de cada uma de suas instâncias (CRUZ, 2010).

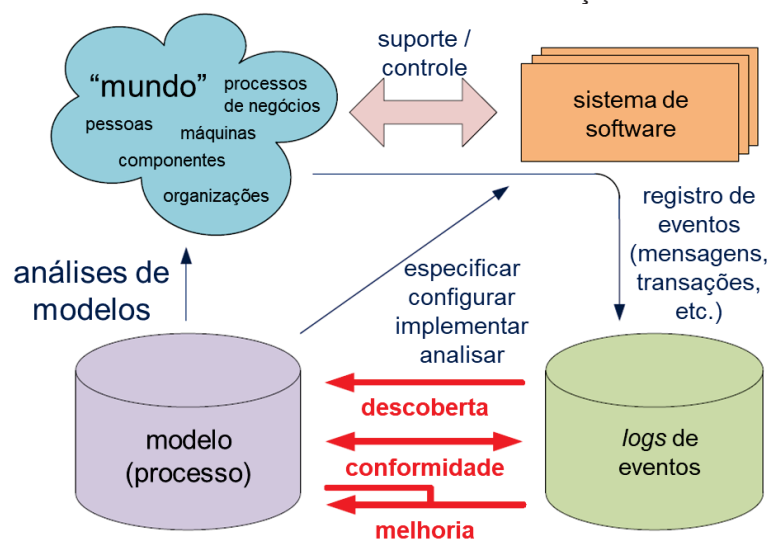
Com isso em mente, ao contrário de produtos, processos são menos tangíveis e, por vezes, podem existir apenas nas mentes das pessoas, dificultando que sejam “materializados”. Por exemplo, é possível que todas as informações presentes em um cadastro de pacientes sejam armazenadas eletronicamente e/ou impressa. Porém, no caso de processos isso torna-se mais complexo. Processos podem ser fruto de comportamentos humanos específicos ou podem ser parcialmente controlados por procedimentos ou sistemas de informação. Associado a isso, diversos processos são documentados na forma de modelos normativos ou descritivos feitos por pessoas e que não refletem, necessariamente, o que ocorre no processo real, resultando em um distanciamento entre modelo e realidade (VAN DER AALST, 2010).

Dessa forma, a mineração de processos apresenta três abordagens complementares para a produção de conhecimento a partir dos registros obtidos (VAN DER AALST, 2016):

- **descoberta** – busca produzir um modelo baseado nos registros de evento sem que haja nenhuma informação obtida *a-priori*;
- **conformidade** – os registros de eventos são confrontados com um modelo pré-existente para verificação de aderências do que efetivamente ocorre com o que é o planejado;
- **melhoria** – busca estender ou melhorar um modelo de processo existente usando informações dos *logs* gravados.

O relacionamento entre os elementos de mineração de processos citados encontra-se resumida na FIGURA 5; os processos são suportados/controlados por sistemas de informação, que armazenam *logs* de eventos e geram um modelo por meio do uso da mineração de processos. Por fim, é feita a análise do modelo e o ciclo recomeça.

FIGURA 5 - ELEMENTOS ENVOLVIDOS NA MINERAÇÃO DE PROCESSOS



FONTE: (VAN DER AALST, 2010 tradução nossa).

Para que seja feita a análise dos *logs* por meio da mineração de processos, existem diversas ferramentas disponíveis no mercado. Cada uma com uma com características próprias e adequadas a contextos específicos. Uma visão geral é apresentada na próxima seção.

### 2.3.1 Ferramentas de mineração de processos

A mineração de processos teve seu início com o uso de algoritmos de máquinas de estados finita (em inglês FSM – *finite-state machines*) e inferência gramatical (COOK; WOLF, 1998). Porém, tais abordagens eram aplicáveis apenas a problemas simples. O primeiro a atacar problemas mais complexos (ex: processos com tarefas duplicadas) foi Herbst (2000), com o uso da busca guiada de verossimilhança de log em modelos de fluxo de trabalho. Atualmente existem diversos algoritmos (ex: cadeias de Markov, algoritmo  $\alpha$ , dentre outros) para que seja feito o sequenciamento dos processos.

Normalmente existem duas formas de apresentação desses algoritmos (1) na forma de um programa independente desenvolvido por um grupo de pesquisadores e (2) a inserção do algoritmo dentro de algum software por meio de *plug-ins*.

Os *softwares* mais utilizados para a mineração de processos são o ProM<sup>12</sup>, que possui uma abordagem mais acadêmica, e o Disco<sup>13</sup>, que possui uma abordagem mais empresarial. Para o desenvolvimento da presente dissertação foi inicialmente adotado o ProM, mas não foi possível processar o volume de dados com a memória disponível na máquina virtual do Java (6gb). Por outro lado, utilizando a segunda opção (Disco), foi possível realizar o processamento dos dados de forma adequada, sendo então o *software* adotado para o desenvolvimento da pesquisa.

O Disco é um software de mineração de processos focado no público empresarial. Sua aplicação baseia-se na utilização de *logs* de eventos dos processos empresariais e trabalha baseado nas informações primárias da mineração de processos: *Case ID*, *Timestamp* e atividade. Sua análise fornece diversas métricas para que seja feita a análise do processo com respeito a tempo total de atravessamento, uso de recursos, análise de conformidade, entre outras (ALVAREZ et al., 2018).

### 2.3.2 Mineração de processos na área da saúde

A aplicação da mineração de processo na área da saúde é um campo que vem apresentando resultados interessantes e que podem ser utilizados na melhoria dos processos decisórios. Tais mudanças impactam diretamente na redução dos custos e na melhoria da qualidade do tratamento (YANG; SU, 2014). A aproximação entre a mineração de processos e a área da saúde teve seu início nos anos 2000, quando diversas instituições de saúde iniciaram a adoção de sistemas de gestão computadorizados que eram capazes de gerar e armazenar os *logs* de eventos das diversas atividades desenvolvidas (RIZ; SANTOS; LOURES, 2016). Desde então, existe uma crescente no que é produzido na literatura (ROJAS et al., 2016).

Os Sistemas de Informação Hospitalar (SIH) armazenam as informações das atividades realizadas pelos recursos hospitalares na forma de *logs* e, quando submetidos à mineração de processos, fornecem informações valiosas na tomada de

---

<sup>12</sup> <http://www.promtools.org>

<sup>13</sup> <https://fluxicon.com/disco/>

decisão. Porém, apesar da evolução dos algoritmos de descoberta de processos e conformidade, quando é discutido o processo de desenvolvimento de um projeto de mineração de processos, ainda não existe a sistematização adequada. Essa falta de uniformidade dificulta que as aplicações de mineração de processos sejam replicáveis (FEI; MESKENS; MOREAU, 2009; FUNKNER; YAKOVLEV; KOVALCHUK, 2017; STAAL, 2010).

Além disso, as aplicações têm sido realizadas utilizando metodologias generalistas. Esse modelo “*one size fits all*” desconsidera as particularidades de cada contexto ao mesmo tempo em que não pode aprofundar-se no detalhamento das etapas para não perder generalidade (BOZKAYA; GABRIELS; WERF, 2009). Como consequência da falta de sistematização da aplicação e da utilização de metodologias generalistas, a implantação torna-se mais lenta e menos precisa, pois exige que o analista de processos passe por toda a curva de conhecimento (GANDULFO, 2016)

Tais dificuldades serão exploradas com mais detalhes na seção 4.2 *Revisão Sistemática da Literatura*.

### 3. METODOLOGIA DA PESQUISA

Nessa seção são apresentados os procedimentos e parâmetros que foram utilizados para o desenvolvimento da pesquisa de modo a atingir os objetivos.

#### 3.1 CARACTERIZAÇÃO DA PESQUISA

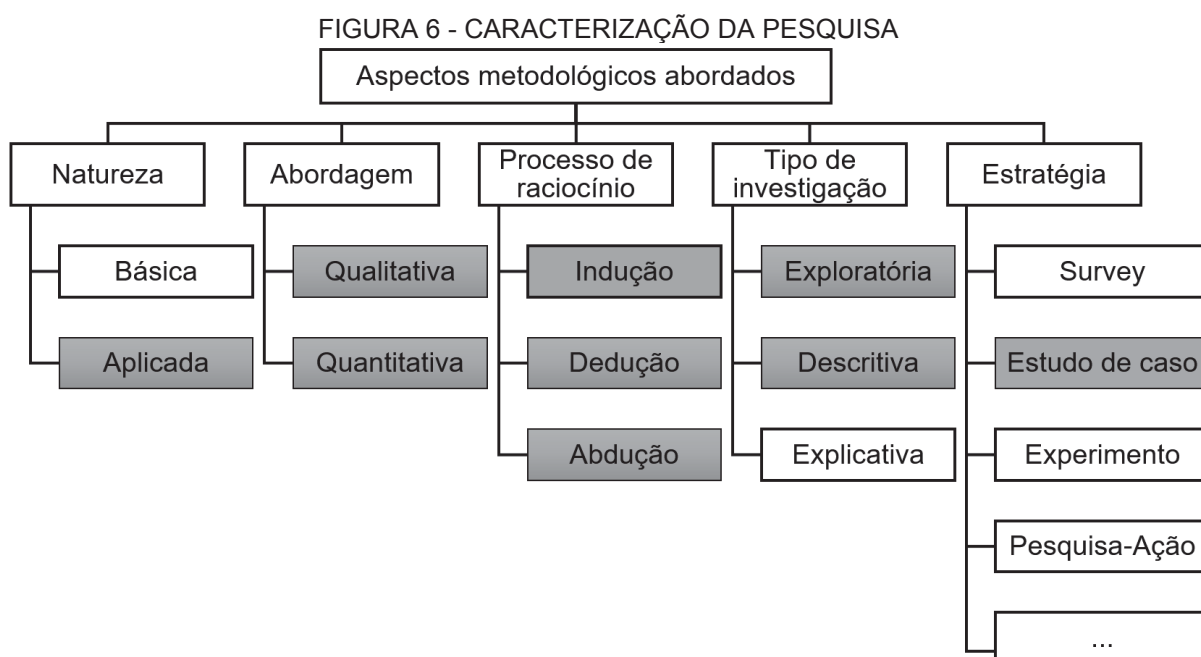
O presente trabalho busca adaptar uma metodologia de aplicação de mineração de processos ao contexto da saúde e, por meio de um estudo de caso na Unidade do Laboratório de Análises Clínicas do de um Hospital-escola terciário, traçar um diagnóstico e sugerir alguns caminhos a serem tomados. Dessa forma, o método de pesquisa utilizado enquadra-se como:

- **Natureza “Aplicada”**: o objetivo é voltado à geração de conhecimentos aplicados à realidade do hospital em questão e dirigidos ao diagnóstico de problemas específicos da Unidade do Laboratório de Análises Clínicas – CHC (SILVA; MENEZES, 2005).
- **Abordagem mista “Qualitativa” e “Quantitativa”**: a pesquisa possui essa classificação mista uma vez que o contexto das ações do laboratório será a fonte direta para coleta de dados, assim como os processos e os respectivos significados são os focos principais de abordagem; configurando a abordagem qualitativa. Complementarmente, no período posterior à análise qualitativa, as informações coletadas são traduzidas em números para sua classificação e análise; configurando a abordagem quantitativa (SILVA; MENEZES, 2005).
- **Processo de raciocínio misto “Abdutivo”, “Dedutivo” e “Indutivo”**: A idealização da abordagem passa por um processo criativo. Em seguida, há a explicação dos eventos desenvolvidos no ULAC-CHC com base em premissas e o estabelecimento de relações de lógica entre elas para obter uma conclusão. Por fim, há a generalização da solução para uma certa classe de problemas (CHALMERS, 1993, p. 23).
- **Tipo de investigação “Exploratório” e “Descritivo”**: a pesquisa em questão divide-se em dois momentos. O primeiro, busca uma visão geral, de tipo aproximativo, acerca dos processos desenvolvidos no laboratório. Esse diagnóstico inicial busca formular hipóteses precisas e operacionalizáveis, esclarecendo o problema e permitindo uma investigação mais ampla. O segundo

momento, por sua vez, apresenta o processo de descrição das características dos fenômenos identificados (GIL, 2008).

- **Estratégia “Estudo de caso”**: a presente pesquisa envolve o estudo profundo e exaustivo dos registros de eventos da ULAC-CHC de maneira que se permita o seu amplo e detalhado conhecimento. Sua caracterização segue o que é proposto por Silva & Menezes (2005).

A FIGURA 6 resume os aspectos metodológicos descritos.

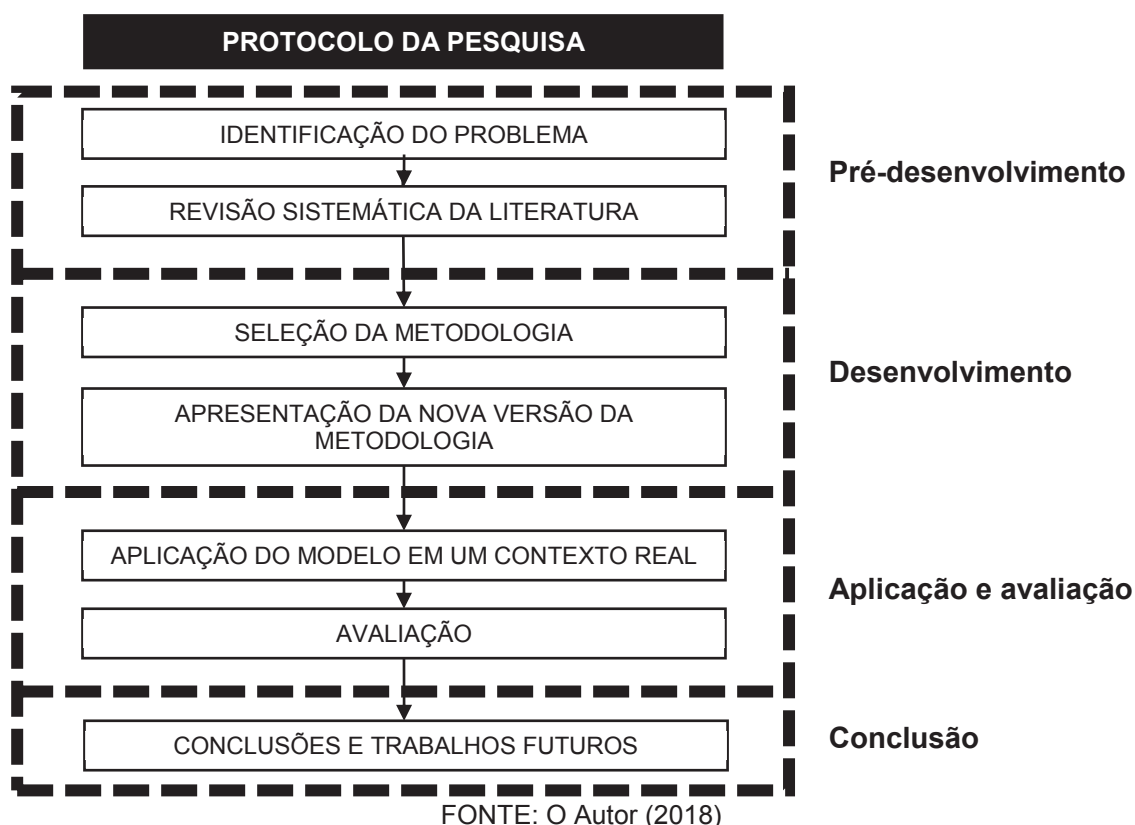


FONTE: O Autor (2017)

### 3.2 PROTOCOLO DE PESQUISA

De modo a atingir os objetivos definidos para a pesquisa, foi elaborado um protocolo para sua condução. O modelo base foi extraído de DRESCH (2013) e teve modificações pontuais para adequar-se ao contexto e aos objetivos propostos. A sequência e métodos adotados estão apresentados na FIGURA 7.

FIGURA 7 - PROTOCOLO DE PESQUISA



### 3.2.1 Pré-desenvolvimento

Nessa etapa inicial da pesquisa foram desenvolvidas ações para que o problema de pesquisa pudesse ser definido e delimitado. A primeira ação dessa etapa foi a “identificação do problema”. Nessa etapa foi necessário justificar porque é importante estudar tal problema. A partir da justificativa de escolha (conforme descrito na introdução), o problema deve ser compreendido de forma clara e objetiva.

Em seguida, há a etapa de “Revisão Sistemática da Literatura” (RSL). A RSL busca proporcionar apoio no processo de explorar problemas similares aos que se está tentando ser solucionado. Ela fornece subsídios para que seja feita a identificação, avaliação e interpretação de toda a pesquisa disponível e relevante para responder uma determinada pergunta de pesquisa, tópico ou fenômeno (KITCHENHAM; CHARTERS, 2007).

Dessa forma, foram realizadas duas RSLs. Os objetivos principais de cada uma das revisões foram: 1) Obtenção de uma visão geral do estado da arte da mineração de processos na área da saúde e identificação de formas de aplicação da mineração

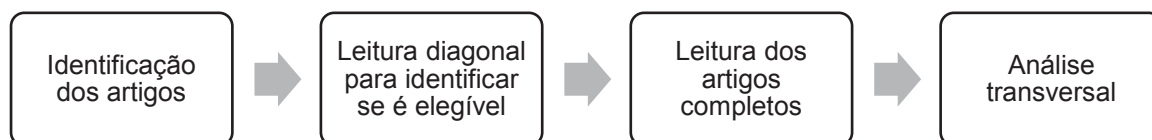


de processos na saúde; 2) Identificação de modelos de mineração de processos no âmbito geral para a seleção do modelo que mais se adeque ao contexto da saúde.

A primeira RSL buscou a obtenção de uma maior intimidade com o tema desenvolvido e a identificação de aplicações da mineração de processos na área da saúde. Dessa forma, foi realizada uma revisão do estudo dos caminhos clínicos na saúde sendo identificadas aplicações da mineração de processos nessa área (4.2.1 *Estado da arte e aplicações da mineração de processos*). A segunda RSL foi realizada para a identificação de metodologias utilizadas no desenvolvimento de projetos de mineração de processos, sem distinção de área de aplicação (4.2.2 *Identificação de modelos de mineração de processos*).

Ambas as RSLs foram desenvolvidas em quatro etapas (FIGURA 8):

FIGURA 8 - ETAPAS PARA AS REVISÕES SISTEMÁTICAS DA LITERATURA



FONTE: (PEREIRA; SANTOS; MACENO, 2018)

- i. **Identificação dos artigos:** foi desenvolvido o protocolo de busca, selecionadas as bases de dados e feita a extração dos artigos.
- ii. **Leitura diagonal para identificar se é elegível:** com os artigos que satisfizeram os critérios de busca iniciais, foi feito o *download* para que pudesse ser realizada uma leitura diagonal (*screening*) buscando a presença das características da RSL em questão.
- iii. **Leitura dos artigos completos:** com os artigos que continham as características desejadas, o texto foi lido em sua íntegra para buscar a compreensão mais profunda da abordagem. Nessa etapa também foram identificados elementos comuns dentre os diversos artigos para que pudessem ser cruzados na próxima etapa
- iv. **Análise transversal:** Por fim, os elementos comuns identificados anteriormente foram utilizados para realizar uma análise de um “corte transversal”, em que os dados são coletados em um ponto no tempo, com base em uma amostra selecionada para descrever uma população nesse determinado momento (RICHARDSON, 1999).

Com o conhecimento do estado da arte, foi possível identificar os elementos que foram utilizados na etapa de desenvolvimento.

### 3.2.2 Desenvolvimento

Nessa etapa, para que seja apresentada a metodologia selecionada, foram associados dois elementos oriundos do pré-desenvolvimento: 1) as metodologias de aplicação da mineração de processos no âmbito geral; 2) comentários, sugestões e percepções apresentados implicitamente nos estudos de caso. Essas informações foram interpretadas com o olhar do contexto de saúde para que pudessem ser consolidadas e apresentadas na forma de uma metodologia de aplicação na área em questão. Após o processo de desenvolvimento, a metodologia foi aplicada e avaliada por meio de um estudo de caso.

### 3.2.3 Aplicação e avaliação

Com a metodologia elaborada, foi realizada a sua aplicação e avaliação em um ambiente real de saúde. O local selecionado foi a Unidade do Laboratório de Análises Clínicas (ULAC) do Complexo Hospital de Clínicas da Universidade Federal do Paraná (CHC-UFPR), que realiza cerca de 500 exames por dia. O CHC-UFPR conta com atendimento 100% SUS e 100% público, onde circulam diariamente cerca de 11 mil pessoas. Possui mais de 200 consultórios e atende cerca de 60 especialidades em seus mais de 400 leitos. Também possui mais de 3.000 funcionários.

### 3.2.4 Conclusão

Por fim, foi realizada a consolidação dos achados e feitas sugestões para trabalhos futuros.

## 4 PRÉ-DESENVOLVIMENTO

O pré-desenvolvimento busca trazer informações pertinentes ao contexto de mineração de projetos dividido em duas partes. A primeira é a identificação do problema, onde é apresentada a lacuna identificada na literatura. Já na segunda, são realizadas revisões sistemáticas referentes a tópicos que sirvam de ferramenta para responder à questão de pesquisa.

### 4.1 IDENTIFICAÇÃO DO PROBLEMA

Na ação de “identificação do problema”, pôde-se estabelecer uma conexão com o que foi descrito na seção de justificativa, em que a área de mineração de processos na saúde carece de normatização e sistematização. Mesmo contemplando diversas ferramentas e algoritmos, não possui uma padronização adequada para sua replicação. Esse fator age negativamente na curva de aprendizado de processos de mineração (dificultando-o). Dessa forma, o problema identificado vai ao encontro dos objetivos de pesquisa e é:

*“Falta de normatização e sistematização da metodologia em processos de mineração de processos e saúde”*

### 4.2 REVISÃO SISTEMÁTICA DA LITERATURA

Conforme mencionado no protocolo (Seção 3.2 *PROTOCOLO DE PESQUISA*), foram realizadas duas RSLs. Cada uma delas com um objetivo específico: 1) Identificação do estado da arte e aplicações da mineração de processos; 2) Identificação de modelos de mineração de processos.

Ambas seguindo a mesma sequência de etapas: i) Identificação dos artigos; ii) Leitura diagonal para identificar se é elegível; iii) Leitura dos artigos completos; iv) Análise transversal.

#### 4.2.1 Estado da arte e aplicações da mineração de processos

Essa RSL inicial buscou a obtenção de uma maior intimidade com o tema desenvolvido e a identificação de aplicações da mineração de processos na área da saúde. Essa seção encontra-se publicada em Pereira, Santos & Maceno (2018).

##### 4.2.1.1 Identificação dos artigos

O protocolo de busca foi elaborado por meio da união de dos protocolos apresentados por revisões sistemáticas já realizadas em áreas correlatas: mineração de processos em oncologia (KURNIATI et al., 2016), mineração de dados e de processos (MAITA et al., 2015) e mineração de processos em saúde (ROJAS et al., 2016). Além da seleção de termos chave das revisões, alguns outros foram adicionados para complementar a chave de busca. Os termos utilizados no protocolo estão apresentados conforme o trabalho de origem das palavras-chave e o grupo a que pertencem: ferramentas, área da saúde e análise, conforme ilustrado pelo QUADRO 1.

QUADRO 1 - AGRUPAMENTO DOS TEMOS PARA A RSL

Ferramenta		Área da saúde		Análise
“mining process” (1) “process mining” (1) (2) (3) “workflow mining” (1) (3) “workflows mining” (1) “mining process” (1) “mining processes” (1) “mining workflow” (1) “mining workflows” (1)	AND	“Healthcare” (3) “Health care” (4) “Hospital” (4) “Medical process” (4) “Healthcare process” (4)	AND	“Patient flow” (2) “Event log” (2) “Clinical pathways” (4) “Pathway analysis” (2) “Pattern mining” (4) “Social network analysis” (4) “Process mapping” (4)
ORIGEM DA PALAVRA-CHAVE <div>             1 Maita et al. (2015)              2 Kurniati, 2016              3 Rojas et al., 2016              4 O autor (2017)           </div>				

FONTE: O Autor (2017)

Consolidando os termos presentes no QUADRO 1 em uma única *string* de busca, obtém-se:

((“mining process” OR “process mining” OR “workflow mining” OR “workflows mining” OR “mining process” OR “mining processes” OR “mining workflow” OR “mining workflows”) AND (“Health care” OR “Hospital” OR “Medical process” OR “Healthcare process”) AND (“Event log” OR “Clinical pathways” OR “Pathway analysis” OR “Pattern mining” OR “Social network analysis” OR “Process mapping”))

A partir dessa *string*, foram realizadas buscas utilizando o *login* institucional da Universidade Federal do Paraná em bases de dados multidisciplinares (1) Scopus, (2) WebOfScience, (3) ScienceDirect, (4) GoogleScholar e em bases de dados especializadas (5) PubMed. Tal escolha das bases foi feita devido à proximidade dos temas presentes em cada uma, que vão ao encontro do tópico da pesquisa: Mineração de processos na área da saúde. As buscas ocorreram em dezembro de 2017. Outra restrição aplicada foi o idioma em que se encontra o manuscrito. Uma vez que o inglês pode ser tomado como a *lingua franca* da ciência<sup>14</sup> (MENEHINI; PACKER, 2007), foram utilizados documentos apresentados apenas nesse idioma.

Foram também feitas restrições com relação ao número de citações do documento em questão, tendo sido utilizados documentos que possuísem pelo menos uma citação. Esse critério foi adotado por dois motivos: (1) permite que os artigos sem citações sejam excluídos, reduzindo o tempo de leitura e (2) concentra-se nos que tiveram, de alguma forma, utilidade para algum pesquisador<sup>15</sup>. Um ponto de ressalva encontra-se na “inércia” presente entre a publicação do documento e sua primeira citação (i.e. documentos muito recentes precisam ser primeiro encontrados para depois serem citados) conforme estudos preliminares apresentados por Riggs et al. (2012).

Por fim, foram utilizados os artigos com data de publicação a partir de 2007 para que a revisão se mantivesse atualizada. Um resumo do protocolo encontra-se apresentado no QUADRO 2.

---

<sup>14</sup> Meneghini & Packer (2007) a descoberta feita por cientistas alemães de uma relação entre o tabagismo e câncer de pulmão no início da década de 1930 e publicada em alemão. Porém, tais resultados foram ignorados até a década de 1960, quando cientistas britânicos e estadunidenses “redescobriram” o *link*. A partir disso, foi que as políticas de saúde pública tiveram a inserção de programas incentivando a redução do uso do cigarro.

<sup>15</sup> A discussão sobre a “velocidade da ciência” é algo que cabe ser mencionado. Com a evolução dos meios de comunicação, o acesso a pesquisas torna-se mais fácil, assim como sua divulgação, e é necessário que critérios de “qualidade” da pesquisa sejam adotados. Tal necessidade ocorre devido à adoção em massa do conceito de “*publish or perish*”, onde existe uma crescente na produção acadêmica em número absoluto, mas não necessariamente em qualidade. Atualmente, o critério mais comumente utilizado é o número de citações que, apesar de cumprir o papel proposto, suscita discussões nos encontros de promoção da ciência acerca da adoção de outras alternativas. Como ilustração dessa reflexão, está a citação do ganhador do prêmio Nobel de física Peter Higgs: “Eu não seria suficientemente produtivo para o sistema acadêmico atual”. <https://www.theguardian.com/science/2013/dec/06/peter-higgs-boson-academic-system>

QUADRO 2 - RESUMO DO PROTOCOLO DA PRIMEIRA RSL

STRING DE BUSCA	((“mining process” OR “process mining” OR “workflow mining” OR “workflows mining” OR “mining process” OR “mining processes” OR “mining workflow” OR “mining workflows”) AND (“Health care” OR “Hospital” OR “Medical process” OR “Healthcare process”) AND (“Event log” OR “Clinical pathways” OR “Pathway analysis” OR “Pattern mining” OR “Social network analysis” OR “Process mapping”))
BASES	(1) Scopus / (2) WebOfScience / (3) Science Direct / (4) GoogleScholar / (5) PubMed
IDIOMA	Inglês
DATA DE PUBLICAÇÃO	>= 2007
PERÍODO DA BUSCA	Dezembro de 2017
LOGIN	Universidade Federal do Paraná
IMPACTO SJR DO PERIÓDICO	>= 1.000
CITAÇÕES	>= 1

FONTE: O Autor (2017)

#### 4.2.1.2 Leitura diagonal para identificar se é elegível

Nessa etapa, foi feito o *download* dos artigos que satisfizeram os critérios de busca iniciais exibidos na etapa anterior (493 artigos) para que pudesse ser realizada uma leitura diagonal (*screening*) buscando a presença de estudos que apresentassem aplicações. Foram abertas duas exceções na elegibilidade dos documentos para dissertações de mestrado (GUPTA, 2007; RAMOS, 2009), devido à sua menção constante nos artigos e à presença de uma versão mais ampla da discussão do processo de aplicação das ferramentas nos processos de saúde.

#### 4.2.1.3 Leitura dos artigos completos

Com os artigos que continham aplicações na saúde, foi possível desenvolver a próxima etapa de Leitura dos artigos completos, onde o texto foi lido em sua íntegra para buscar a compreensão mais profunda da abordagem. Conforme cada texto era lido, as percepções e observações do autor do artigo foram sendo planilhadas. Após a tabulação de todas as informações extraídas de todos os artigos, foram levantados os elementos comuns presentes em cada um deles. Foi então identificado que cinco elementos estavam presentes de forma recorrente nos artigos, e foram então utilizados na próxima etapa: 1) clusterização; 2) filtragem; 3) *Key Performance*

*Indicators*<sup>16</sup> (KPIs); 4) aproximação com os *stakeholders*<sup>17</sup> e 5) reconhecimento da importância do uso de Representações Gráficas.

A **clusterização**, que também foi mencionada como “granularidade”, é o agrupamento de elementos similares para serem analisados em conjunto. Isso significa medir o grau de similaridade dos casos e utilizar essa informação para dividir o conjunto de casos em subgrupos mais homogêneos. Não adotar essa prática pode resultar em um *log* que contenha muitas atividades distintas que resultaria em um modelo excessivamente detalhado (modelo de espaguete), que é difícil de interpretar (MANS et al., 2009).

A “**filtragem**” busca lidar com a informação de forma que ela seja apresentada na quantidade adequada. Tal prática evita, por exemplo, que eventos que não apresentam a frequência de ocorrência parelha com os demais sejam removidos da análise. Também existem sugestões de que seja criado um *log* separado para esses eventos diferentes (GÜNTHER et al., 2008).

A presença de “**KPIs**” também apareceu mencionada como “estatísticas do modelo”. Dentre as diversas opções que podem ser adotadas estão: tempo total de atravessamento, tempos de espera, número de consultas que se repetem para um mesmo paciente, etc. Esses indicadores são utilizados, em especial, para a comparação entre o que foi planejado e o que efetivamente ocorreu ou para verificar o efeito de alguma alteração no modelo.

A aproximação do analista de mineração de processos com os **stakeholders** (tomadores de decisão) foi mencionada em diversos trechos. O objetivo de tal aproximação é fazer com que haja o alinhamento das expectativas e o entendimento dos resultados de forma mais simples para os especialistas envolvidos (médicos, gestores, etc.). Essa aproximação busca auxiliar também o analista a compreender o processo de modo mais holístico para poder identificar quais elementos são passíveis de clusterização e filtragem, por exemplo.

Por fim, o reconhecimento da importância de **representações gráficas** também é um ponto em que é possível comparar os artigos. Sua utilidade está na facilitação da comunicação interdisciplinar com os *stakeholders*. Essas representações são úteis tanto antes da mineração propriamente dita (etapa de

---

<sup>16</sup> Indicadores chave de desempenho

<sup>17</sup> Partes interessadas, tomadores de decisão

entendimento do processo) quanto depois (apresentação dos resultados). Com as dimensões definidas, foi possível proceder para a etapa de análise transversal.

#### 4.2.1.4 Análise transversal

Por fim, a etapa de Análise transversal utilizou os elementos comuns identificados anteriormente para realizar uma análise de um “corte transversal”. Foi realizada a listagem dos estudos de caso apresentados com os elementos comuns listados conforme a presença (✓) ou ausência (-). O resultado encontra-se no QUADRO 3.

Por meio da análise do QUADRO 3 é possível notar que as cinco dimensões selecionadas estão presentes consistentemente nos documentos, com destaque para a clusterização (82%) e a utilização de KPIs (76%). Também foi possível identificar que todos os casos apresentam mais do que uma opção, indicando que pouco provavelmente um projeto de mineração de processos conseguirá ser desenvolvido satisfatoriamente sem que essa combinação de elementos seja feita.

QUADRO 3 - APLICAÇÕES DA MINERAÇÃO DE PROCESSOS NA ÁREA DA SAÚDE

REFERÊNCIA	Estudo de caso	C	F	KPIs	S	RG
(GUPTA, 2007)	Unidade de Tratamento Intensivo (UTI) + Derrame	✓	-	✓	-	-
(MANS et al., 2009)	Derrame + Oncologia ginecológica	-	✓	-	✓	-
(BLUM et al., 2008)	Laparoscopia	✓	-	✓	-	✓
(GÜNTHER et al., 2008)	Máquinas de raio-X	✓	-	✓	-	-
(LANG et al., 2008)	Fluxo em equipamentos radiológicos	✓	-	✓	-	-
(RAMOS, 2009)	Oncologia ginecológica	✓	✓	✓	✓	✓
(RIEMERS, 2009)	Pé diabético e mamografia	✓	✓	✓	✓	✓
(POELMANS et al., 2010)	Câncer de mama	✓	-	✓	-	✓
(BAKER et al., 2017)	Câncer de mama + Câncer colorretal	✓	-	✓	-	-
(PERIMAL-LEWIS et al., 2012)	Fluxo de pacientes na enfermagem	✓	✓	✓	✓	-
(REBUGE; FERREIRA, 2012)	Departamento de emergência	✓	✓	✓	✓	✓
(WANG et al., 2013)	Pacientes com diabetes	✓	-	-	-	✓
(DEFOSSEZ et al., 2014)	Câncer de mama unilateral não metastático	✓	-	✓	-	-
(BETTENCOURT-SILVA et al., 2015)	Câncer de próstata	-	✓	✓	✓	✓
(DELIAS et al., 2015)	Departamento de emergência	✓	-	✓	✓	-
(HUANG et al., 2015)	Angina instável	✓	-	-	✓	✓
(CHO et al., 2017)	Centro de neurociência	-	✓	-	✓	✓

\*Clusterização (C), Filtragem (F), KPIs, Stakeholders (S) e Representações Gráficas (RG)

FONTE: O Autor (2018)



Além do apresentado no QUADRO 3, outros comentários relevantes foram identificados ao longo da leitura completa dos trabalhos e encontram-se enumerados a seguir:

1. *Logs* com informações gravadas em um intervalo diário, ao invés de um *timestamp* (dd-MM-aaaa HH:mm) apresentam dificuldade no ordenamento das atividades. Ex: atividades que ocorreram no início da manhã serão interpretadas da mesma forma que aquele que ocorreram no final da tarde. (MANS et al., 2009).
2. Outra alternativa para lidar com os modelos do tipo “espaguete” é quebrar o *log* em dois ou mais sub-logs até que estes se tornem simples o suficiente para serem analisados claramente (MANS et al., 2009).
3. Um dos melhores sistemas para se extrair os dados é o sistema de faturamento/cobrança, uma vez que o registro de todas as atividades desenvolvidas por cada paciente deve ser confiável; pois deve ser cobrado o valor correto de cada paciente (RAMOS, 2009).
4. Recomenda-se que um *log* contenha: ID do paciente, nome da atividade, departamento envolvido, *timestamp* e, opcionalmente, atributos para desenvolver análises adicionais (como idade do paciente, por exemplo) (RAMOS, 2009).
5. Uma vez que os *logs* são compostos por dados sensíveis (informações de pacientes, cobrança, etc.), deve ser tomado especial cuidado com os dados. Sugere criptografá-los (RIEMERS, 2009).

A partir dessa análise, foi realizada a busca por modelos de aplicação de mineração de processos na saúde para que as informações pudessem ser combinadas.

#### 4.2.2 Identificação de modelos de mineração de processos

Dando continuidade às revisões, foi realizada uma segunda busca para a identificação de metodologias utilizadas no desenvolvimento de projetos de mineração de processos, sem distinção de área de aplicação. Foram seguidos os mesmos quatro passos utilizados na revisão da seção 4.2.1 *Estado da arte e aplicações da mineração de processos*.

#### 4.2.2.1 Identificação dos artigos

O protocolo da Identificação dos artigos dessa segunda revisão sofreu alterações devido a algumas percepções obtidas na revisão anterior: 1) a base GoogleScholar foi eliminada pois a extração de artigos não é tão simples como nas outras bases (com arquivos .csv ou .txt) e os resultados não são satisfatórios (apesar de ter sido feito o uso do software *Publish or Perish*<sup>18</sup>, os resultados não apresentaram a mesma qualidade dos artigos extraídos de outras bases com relação aos campos extraídos e ao seu preenchimento); 2) Foram realizadas algumas rodadas de busca com diversas combinações de parâmetros nas bases e, ao contrário do que ocorreu na seção 4.2.1 Estado da arte e aplicações da mineração de processos, não houve uma quantidade de artigos suficiente para desenvolver uma análise satisfatória (apenas 38 artigos). Assim, foram realizadas alterações com relação à origem dos artigos (para permitir que artigos de congressos também pudessem ser incorporados à análise) e com relação às citações (desconsiderando esse critério).

Com relação à chave de busca, foram utilizados termos amplos de busca de metodologias/métodos e implantação de projetos de mineração de processos. Dessa forma, foi feita a busca de artigos em inglês em quatro bases de dados: Scopus, WebOfScience, Science Direct e PubMed com a seguinte chave de busca:

((*"Process mining"*) AND ((*"process mining methodology"*) OR (*"process mining method"*) OR (*"project methodology"*) OR (*"project model"*) OR (*"project deployment"*))))

O resumo do protocolo encontra-se apresentado no QUADRO 4.

QUADRO 4 - RESUMO DO PROTOCOLO DA SEGUNDA RSL

<b>STRING DE BUSCA</b>	(( <i>"Process mining"</i> ) AND (( <i>"process mining methodology"</i> ) OR ( <i>"process mining method"</i> ) OR ( <i>"project methodology"</i> ) OR ( <i>"project model"</i> ) OR ( <i>"project deployment"</i> ))))
BASES	(1) Scopus / (2) WebOfScience / (3) Science Direct / (4) PubMed
IDIOMA	Inglês
DATA DE PUBLICAÇÃO	>= 2007
PERÍODO DA BUSCA	Junho de 2018
LOGIN	Universidade Federal do Paraná

FONTE: O Autor (2018)

<sup>18</sup> <https://harzing.com/resources/publish-or-perish>

#### 4.2.2.2 Leitura diagonal para identificar se é elegível

Com os artigos que satisfizeram os critérios de busca iniciais, iniciou-se a etapa de Leitura diagonal para identificar se é elegível. Da mesma forma como anteriormente, foi feito o *download* dos artigos exibidos na etapa anterior para que pudesse ser realizada uma leitura diagonal (*screening*). O critério, dessa vez, foi a presença da descrição do processo de aplicação de mineração de processos, sem distinção de área. Essa descrição deveria ser algo diferente de simplesmente dizer “a mineração de processos foi aplicada e esses são os resultados”. Foram priorizados, quando possível, os artigos que mencionaram a sequência de etapas presentes desde o momento da concepção do projeto até o seu término.

#### 4.2.2.3 Leitura dos artigos completos

Com os artigos que cumpriam com os requisitos mencionados na leitura diagonal, foi desenvolvida a etapa de Leitura dos artigos completos, onde os modelos identificados na etapa anterior foram explorados mais a fundo. Da mesma forma como feito na revisão anterior (Seção 4.2.1 Estado da arte e aplicações da mineração de processos), foram identificados elementos comuns dentre os diversos artigos para serem utilizados na próxima etapa: 1) Área de aplicação; 2) Etapas do ciclo PDCA; 3) Ênfase; 4) “É iterativo?”. Os elementos (1) e (3) têm o objetivo de categorizar a aplicação. O elemento (2) foi inserido pelo caráter de melhoria dos processos organizacionais (KURNIATI et al., 2016; MAITA et al., 2015; RIGGS et al., 2012). E os elementos (3) e (4) resgata as observações mencionadas na justificativa do presente trabalho.

A “**área de aplicação**” ilustra o espectro de aplicação da mineração de processos no âmbito geral. Além disso, permite compreender as aplicações nas quais há a presença de metodologias mais estruturadas.

As “**etapas do ciclo PDCA**”<sup>19</sup> são uma forma de interpretação do modelo apresentado dentro dos moldes do conceito de Planejar, Desenvolver, Conferir e Agir. Essa dimensão de análise traz informações referentes a quais etapas do processo o modelo está inserido – o foco é maior no início, meio ou fim?

---

<sup>19</sup> Plan, Do, Check, Action

Com relação à “ênfase”, pela leitura dos artigos foram identificadas três possíveis opções: 1) Processo: a ênfase do modelo encontra-se no processo como um todo, suas etapas e como desenvolvê-las; 2) Algoritmo: a ênfase encontra-se no desenvolvimento de algum algoritmo; 3) Pessoas: a ênfase está no uso do conhecimento das pessoas envolvidas (conhecimento tácito) para tornar o modelo do processo mais preciso

Por fim, foi feita a classificação sobre a metodologia ser ou não “**iterativa**”. Essa dimensão foi inserida para que o próprio modelo incentive e permita a correção de erros cometidos ao longo do processo.

#### 4.2.2.4 Análise transversal

Nessa etapa final, a análise transversal, estão apresentadas as metodologias<sup>20</sup> adotadas pelos autores para o desenvolvimento de suas aplicações (FIGURA 10 a FIGURA 9). As figuras estão dispostas listadas por ordem de data de criação da metodologia.

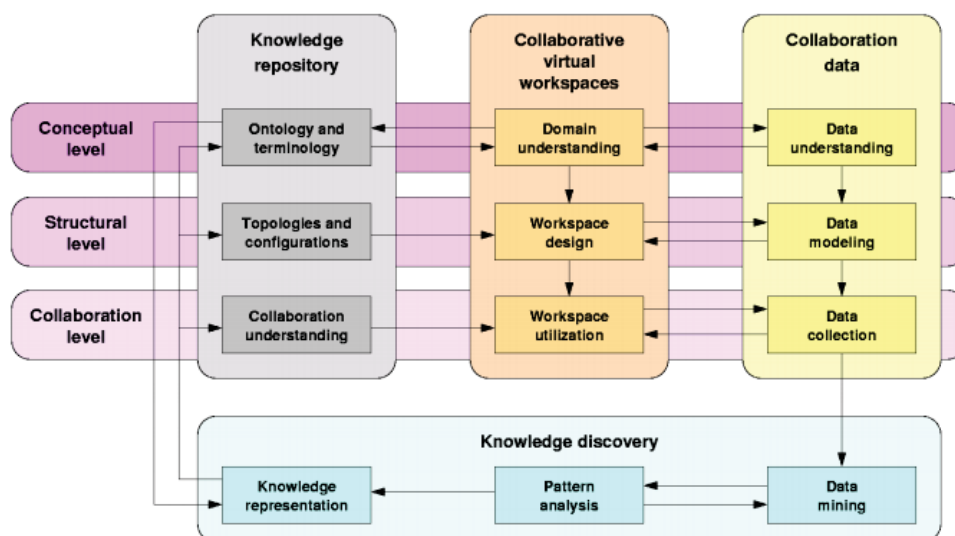
Para a apresentação das metodologias, é possível uma interpretação associada ao período em que cada uma foi desenvolvida. A metodologia de BIUK-AGHAI & SIMOFF (2001) (FIGURA 10) é essencialmente um modelo de mineração de dados. Porém, com sua aplicação em ambientes colaborativos virtuais possibilitou o levantamento de elementos que seriam utilizados na mineração de processos (tais como atividades e recursos).

No ano de 2003 ocorreu a primeira pesquisa formalizada de questões e abordagens em mineração de processos, tornando-se mais madura e com mais ferramentas (van der AALST, 2016). É possível ver que, a metodologia de BOZKAYA, GABRIELS & WERF (2009) (FIGURA 11) já apresenta a inserção da ideia de fluxo, seguida pela metodologia de FANG et al. (2011) (FIGURA 12) que menciona processos explicitamente.

---

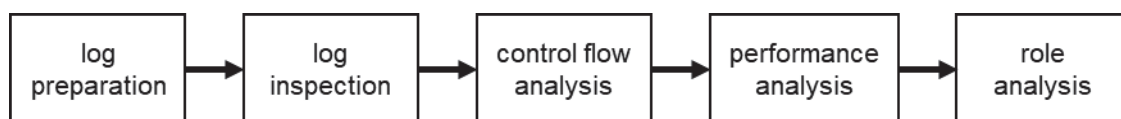
<sup>20</sup> As metodologias estão apresentadas na língua original (inglês) para que não ocorra a descaracterização do modelo devido à tradução

FIGURA 10 - METODOLOGIA DE BIUK-AGHAI &amp; SIMOFF



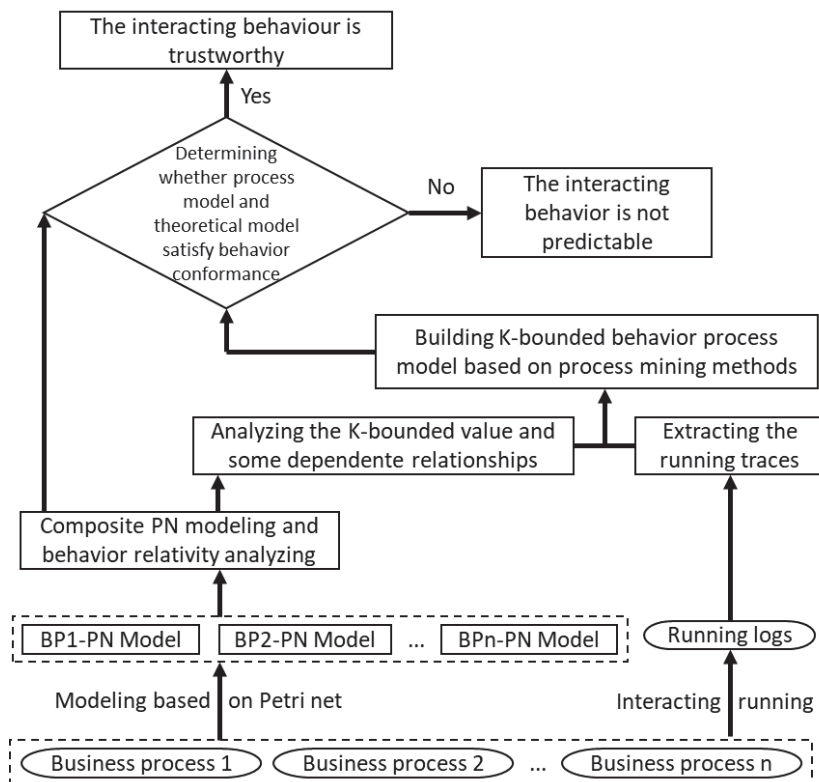
FONTE: (BIUK-AGHAI; SIMOFF, 2001)

FIGURA 11 - METODOLOGIA DE BOZKAYA, GABRIELS &amp; WERF



FONTE: (BOZKAYA; GABRIELS; WERF, 2009)

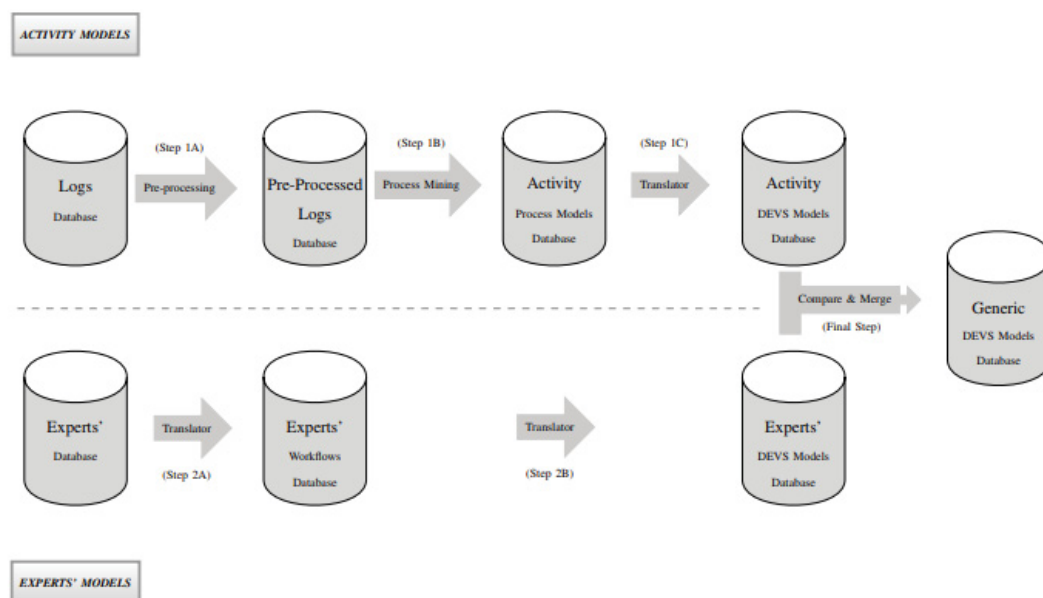
FIGURA 12 - METODOLOGIA DE FANG ET AL.



FONTE: (FANG et al., 2011)

A metodologia de VIALE; FRYDMAN & PINATON (2011) (FIGURA 13) acrescenta dois elementos importantes para um projeto de mineração de processos: modelos de processo e a integração de “*experts*” no processo.

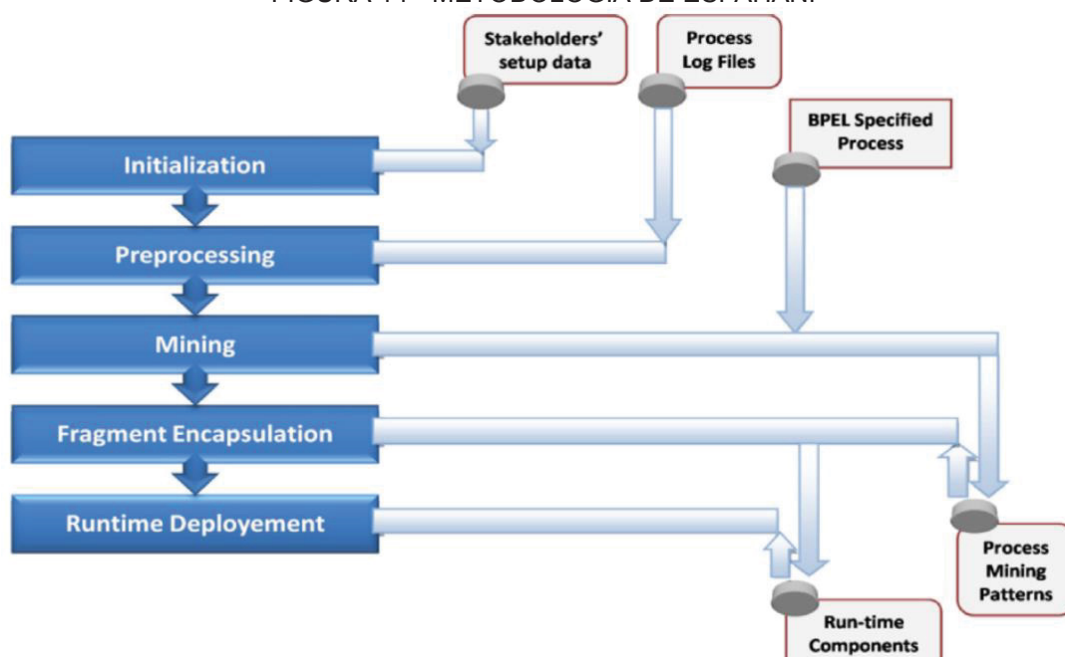
FIGURA 13 - METODOLOGIA DE VIALE, FRYDMAN & PINATON



FONTE: (VIALE; FRYDMAN; PINATON, 2011)

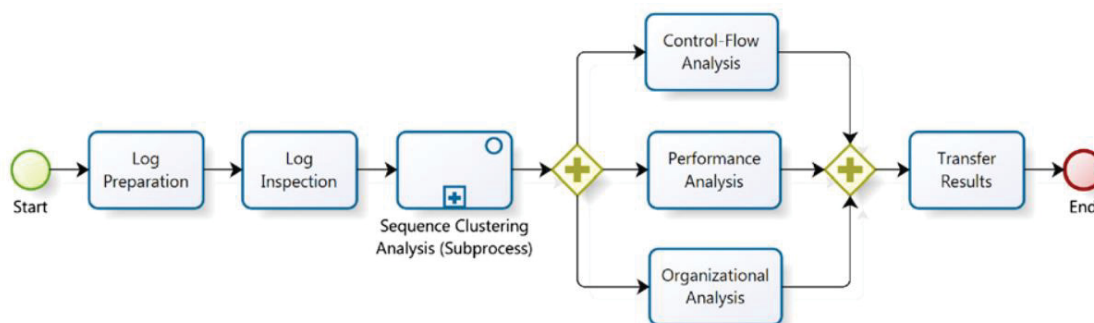
Esfahani et al. (2011) (FIGURA 14) trazem um novo elemento muito importante para um projeto (independentemente do tipo): a presença dos tomadores de decisão (*stakeholders*). Na FIGURA 15, Rebuge & Ferreira (2012) colocam a importância da Clusterização em sua metodologia.

FIGURA 14 - METODOLOGIA DE ESFAHANI



FONTE: (ESFAHANI et al., 2011)

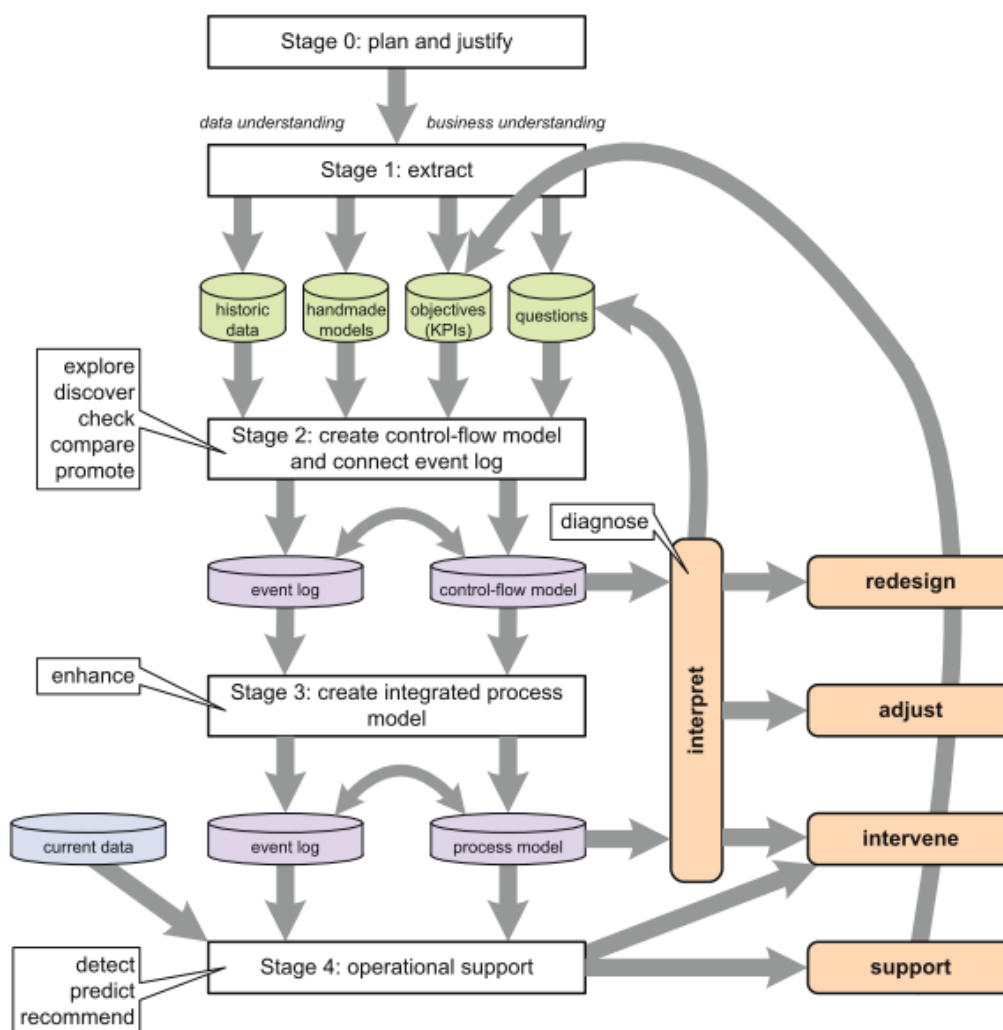
FIGURA 15 - METODOLOGIA DE REBUGE &amp; FERREIRA



FONTE: (REBUGE; FERREIRA, 2012)

Em 2013, van der AALST traz um modelo mais robusto e voltado especificamente para mineração de processos, o *L\* lifecycle model* (FIGURA 16), que incorpora etapas iterativas e uma segmentação mais estruturada dos estágios.

FIGURA 16 – METODOLOGIA DE VAN DER AALST



FONTE: (VAN DER AALST, 2013a)



A metodologia de LEE et al. (2013) (FIGURA 17) volta a enfatizar a importância da clusterização em um modelo mais simplificado. WEERDT (2013) (FIGURA 18) e FERNANDEZ-LLATAS et al. (2015) (FIGURA 20) acrescentam uma etapa de melhoria do processo e SHUKLA, KEAST & CEGLAIREK (2014) (FIGURA 19) coloca a importância da troca de conhecimento por meio de debates em grupo e *brainstormings*.

FIGURA 17 - METODOLOGIA DE LEE ET AL.

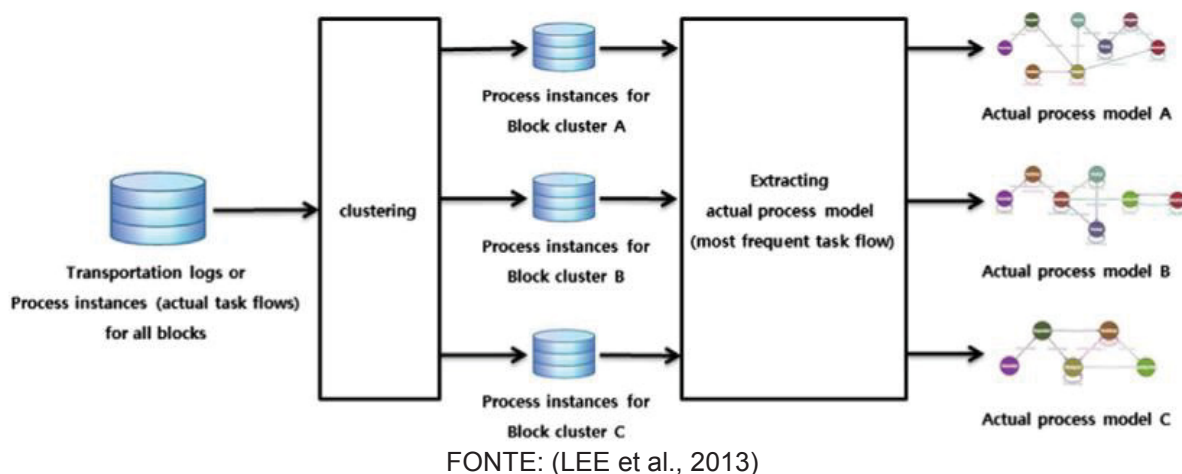


FIGURA 18 - METODOLOGIA DE WEERDT ET AL.

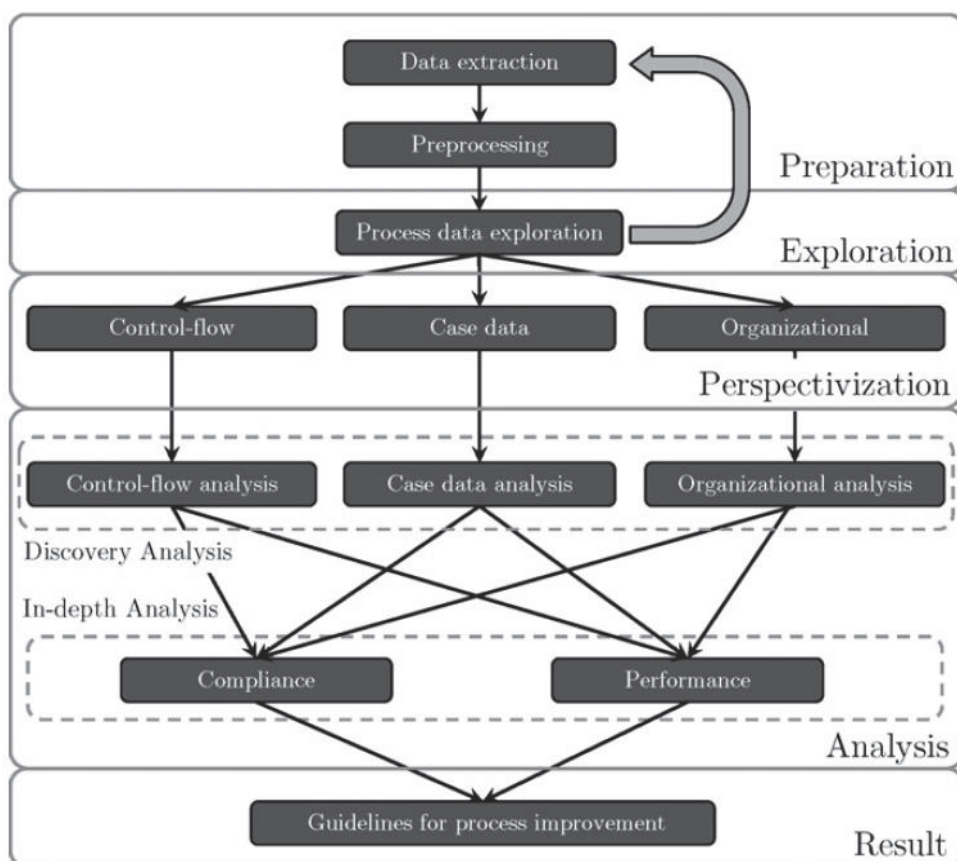
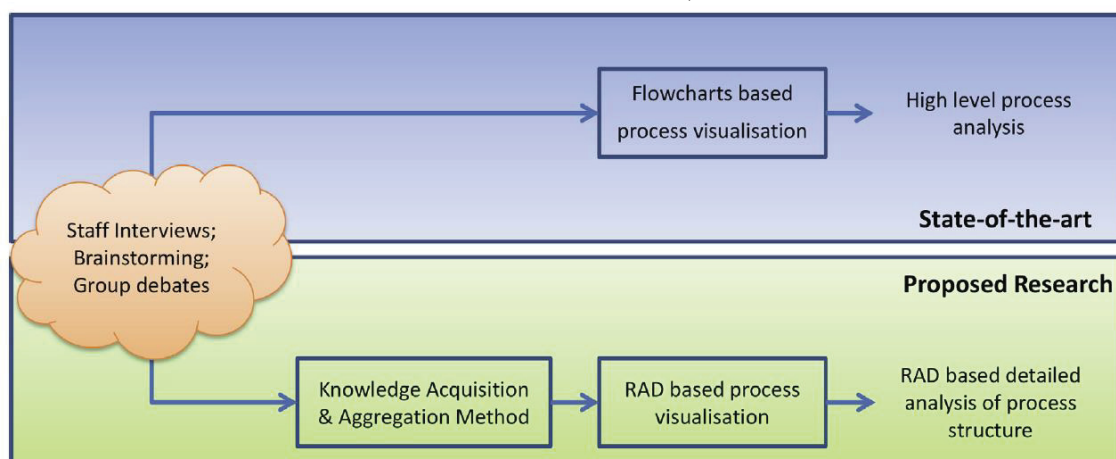


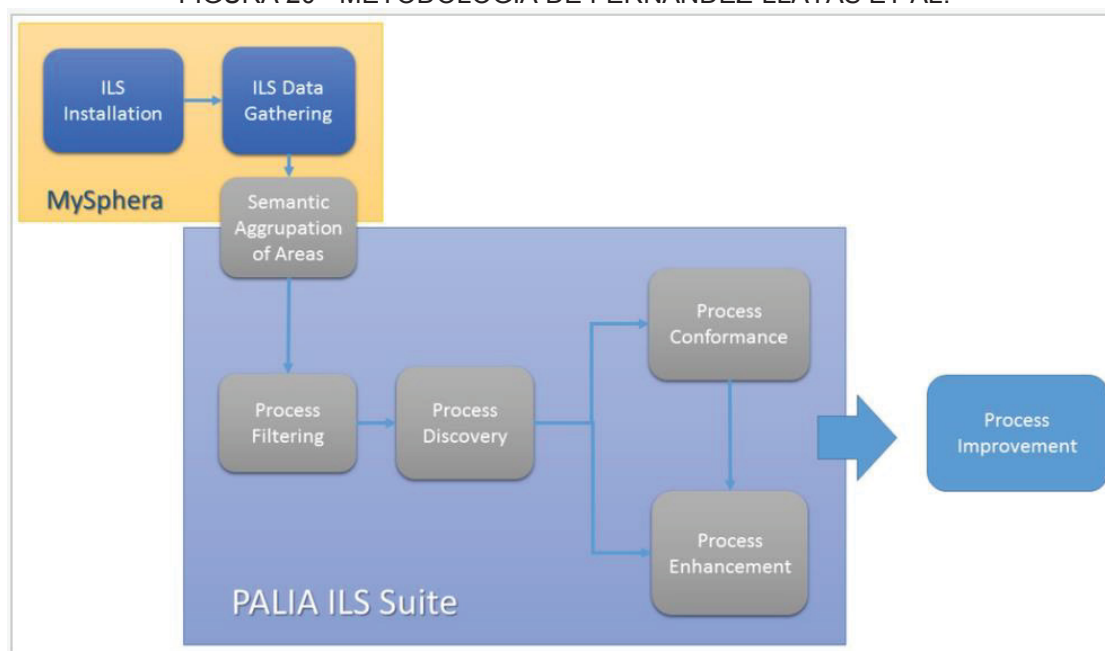


FIGURA 19 - METODOLOGIA DE SHUKLA, KEAST &amp; CEGLAREK



FONTE: (SHUKLA; KEAST; CEGLAREK, 2014)

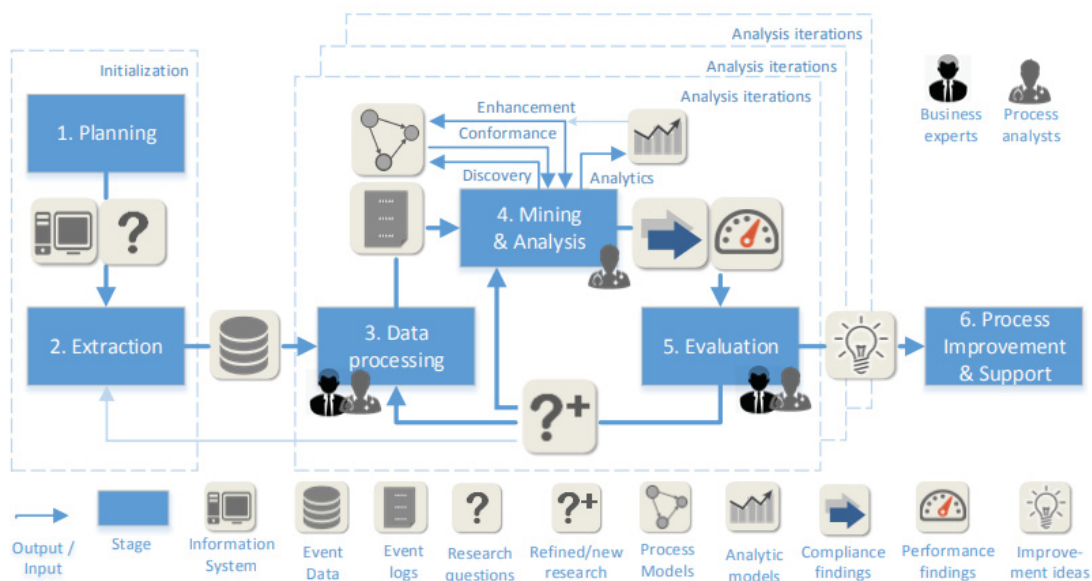
FIGURA 20 - METODOLOGIA DE FERNANDEZ-LLATAS ET AL.



FONTE: (FERNANDEZ-LLATAS et al., 2015)

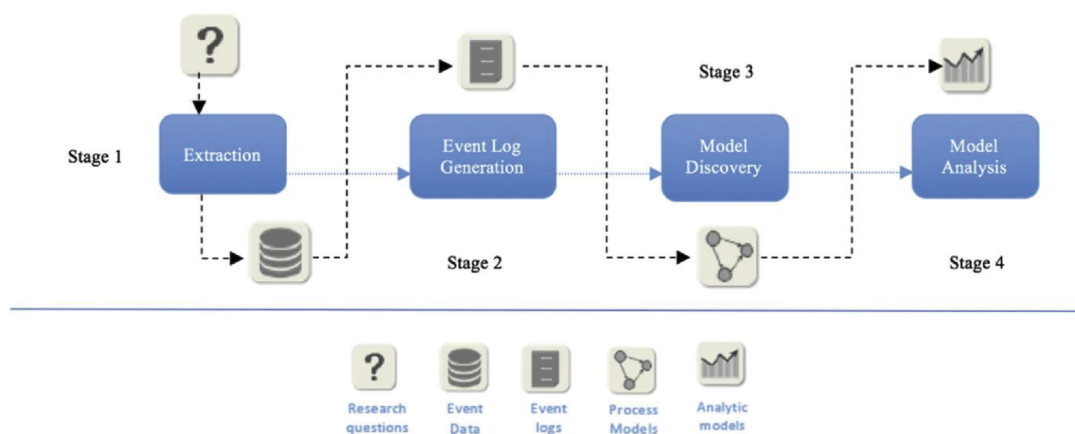
Em 2015, van ECK et al. (FIGURA 21) compõem um modelo que contempla desde as etapas iniciais de planejamento, passando pelo processamento e mineração propriamente dita, terminando na melhoria do processo. Além disso, incorporam os profissionais responsáveis pelo processo no desenvolvimento do projeto. As metodologias de MALDONADO-MAHAUAD et al. (2018) (FIGURA 22) e de EPURE et al. (2016) (FIGURA 23) são versões simplificadas do modelo de van ECK.

FIGURA 21 - METODOLOGIA DE VAN ECK ET AL.



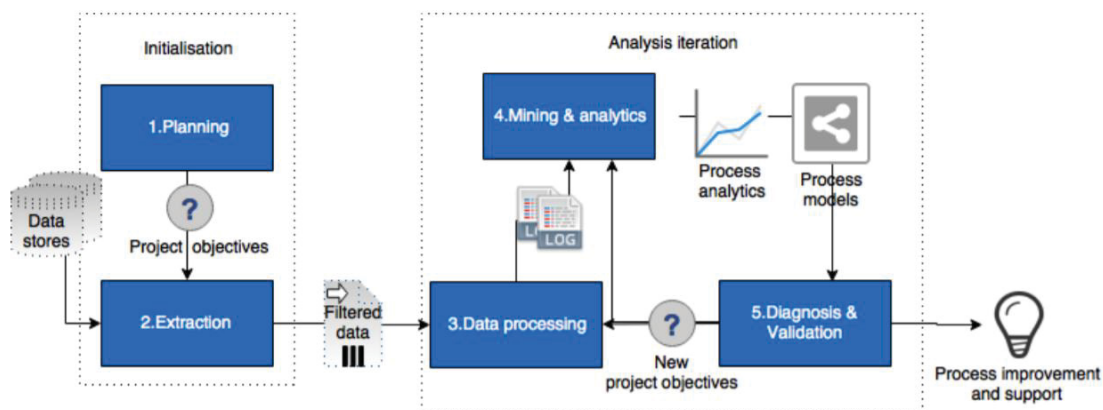
FONTE: (VAN ECK et al., 2015)

FIGURA 22 - METODOLOGIA DE MALDONADO-MAHAUAD ET AL.



FONTE: (MALDONADO-MAHAUAD et al., 2018)

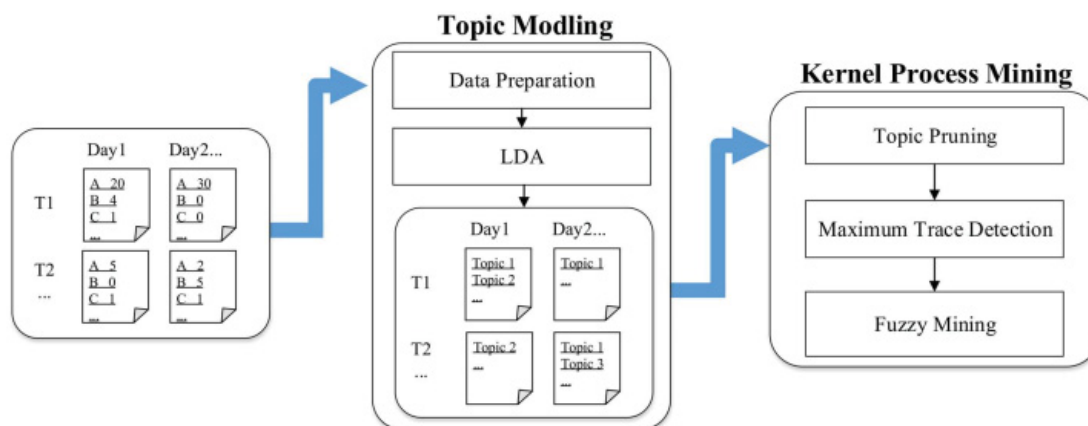
FIGURA 23 - METODOLOGIA DE EPURE ET AL.



FONTE: (EPURE et al., 2016)

A metodologia de XU et al. (2016) (FIGURA 24) é a aplicação de um algoritmo utilizado na mineração de processos.

FIGURA 24 - METODOLOGIA DE XU ET AL.



FONTE: (XU et al., 2016)

LEHTO, HINKKA & HOLLMEN (2017) adicionam mais explicitamente a ideia do uso de KPIs para mensurar o efeito da melhoria, além de apresentar os resultados para as pessoas envolvidas (FIGURA 25).

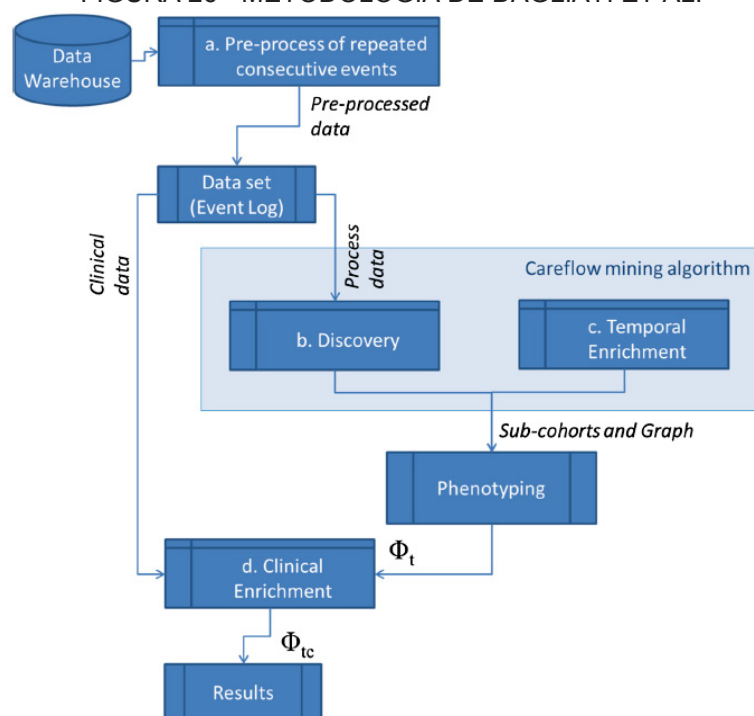
FIGURA 25 - METODOLOGIA DE LEHTO, HINKKA & HOLLMEN

1. Identify the relevant business process and define the case
2. Collect event and case attribute information
3. Create new categorization dimensions
4. Forma a binary classification of cases such that each case is either problematic or successful
5. Select a corresponding interestingness measure based on the desired level of business process improvement effect
6. Find the best categorization rules and attributes
7. Present the results to business people

FONTE: (LEHTO; HINKKA; HOLLMEN, 2017)

A metodologia de DAGLIATI et al. (2017) (FIGURA 26) é totalmente direcionada ao estudo de caso selecionado (pacientes com câncer de mama) e apresenta a ideia de “enriquecimento” dos *logs*.

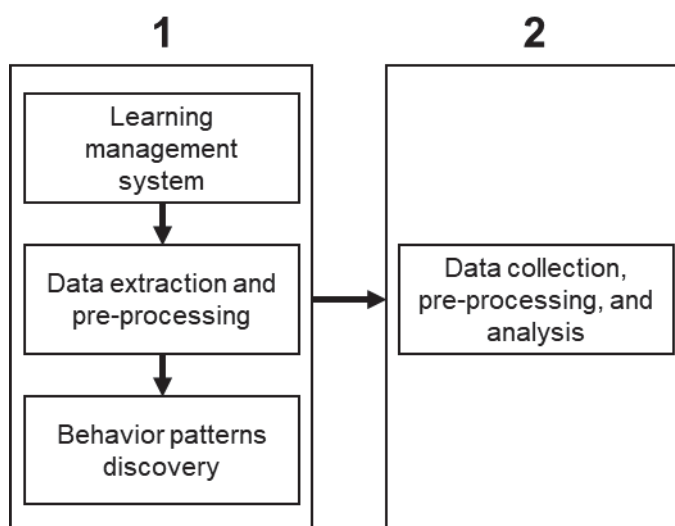
FIGURA 26 - METODOLOGIA DE DAGLIATI ET AL.



FONTE: (DAGLIATI et al., 2017)

A metodologia de JUHANAK, ZOUNEK & ROHLÍKOVÁ (2017) (FIGURA 27) também é direcionada para uma aplicação específica (sistemas de gerenciamento de aprendizado) e foi descrita em forma de texto, sendo esse o motivo de sua adaptação para a forma de figura.

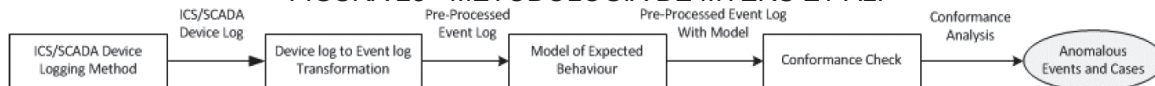
FIGURA 27 - METODOLOGIA JUHANAK, ZOUNEK &amp; ROHLÍKOVÁ



FONTE: Adaptado de (JUHANAK; ZOUNEK; ROHLÍKOVÁ, 2017)

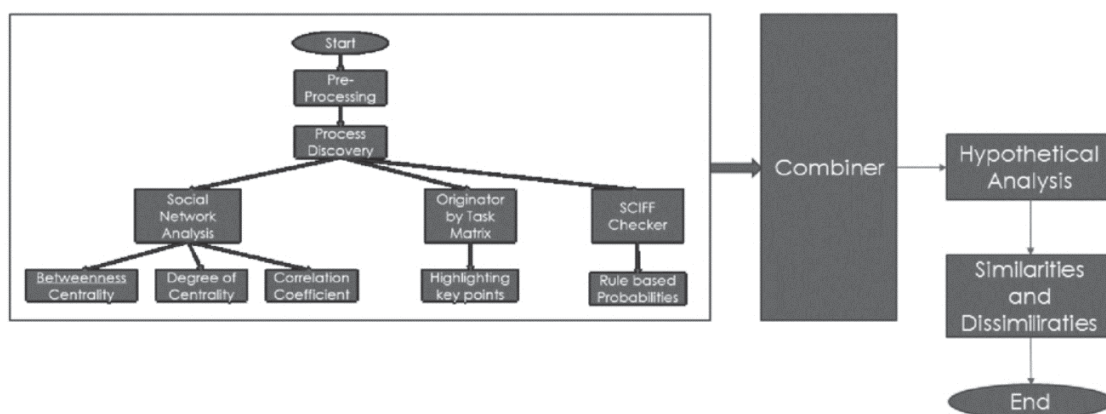
A metodologia de MYERS et al. (2018) (FIGURA 28) e a de KHOWAJA (2018) (FIGURA 29) trazem o uso de verificações de conformidade e análise de rede social, respectivamente.

FIGURA 28 - METODOLOGIA DE MYERS ET AL.



FONTE: (MYERS et al., 2018)

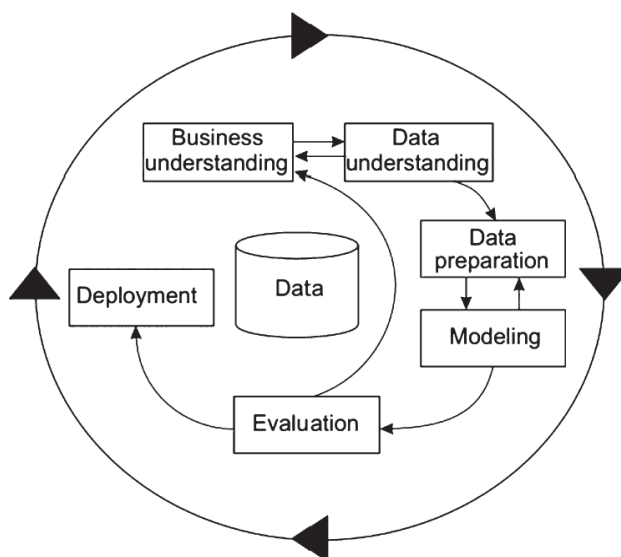
FIGURA 29 - METODOLOGIA DE KHOWAJA



FONTE: (KHOWAJA, 2018)

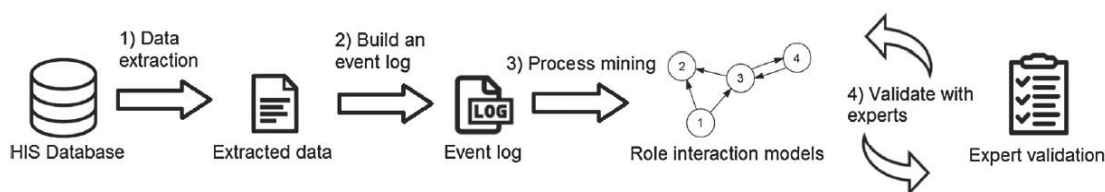
GERHARDT, VALIATI & SANTOS (2018) (FIGURA 30) enfatizam a compreensão tanto dos dados e do processo antes do início da utilização efetiva dos dados, além de elementos iterativos. A presença da iteratividade também é citada na metodologia de ALVAREZ et al. (2018) (FIGURA 31), porém, restringe-se à etapa de validação com *experts*.

FIGURA 30 - METODOLOGIA DE GERHARDT, VALIATI & SANTOS



FONTE: (GERHARDT; VALIATI; SANTOS, 2018)

FIGURA 31 - METODOLOGIA DE ALVAREZ ET AL.



FONTE: (ALVAREZ et al., 2018)

Com as metodologias apresentadas, é possível proceder com a consolidação dos resultados, que se encontra no QUADRO 5, seguida da discussão.

QUADRO 5 – MODELOS DE MINERAÇÃO DE PROCESSO PRESENTES NA LITERATURA: CONSOLIDAÇÃO DAS METODOLOGIAS 1 A 22

#	Aplicação	ETAPAS presentes	Ênfase	Iterativo?
1	Área de trabalho virtual	P, D, C, A	Processo/Algoritmo	Não
2	Saúde	D, C	Algoritmo/Processo	Não
3	Geral	P, D, C	Processo/Algoritmo	Não
4	Microeletrônicos	P, D, C, A	Pessoas / Processo	Não
5	Geral	D, A	Processo/Algoritmo	Não
6	Saúde	D, C	Processo/Algoritmo	Não
7	Geral	P, D, C, A	Pessoas / Processo	Sim
8	Transporte	D, C	Pessoas / Processo	Não
9	Serviços financeiros	P, D, C, A	Pessoas / Processo	Sim
10	Saúde	P, D, C, A	Pessoas / Processo	Não
11	Saúde	D, C, A	Pessoas / Processo	Não
12	Compras	P, D, C, A	Pessoas / Processo	Sim
13	MOOCs*	D, C	Processo	Não
14	Notícias	P, D, C, A	Pessoas / Processo	Sim
15	Saúde	P, D	Algoritmo	Não
16	Banco	P, D, C	Processo/Algoritmo	Não
17	Saúde	D, C	Processo/Algoritmo/Pessoas(pouco)	Não
18	Sistemas de aprendizado	D, C	Processo	Não
19	Detecção de anomalias	D, C	Processo	Não
20	Gerenciamento do tempo	D	Processo/Algoritmo	Não
21	Processo de reembolsos	P, D, C, A	Processo	Sim
22	Saúde	D, C	Pessoas / Processo	Sim

\* MASSIVE OPEN ONLINE COURSE.

FONTE: O Autor (2018)

Por meio do QUADRO 5 é possível notar que, mesmo não tendo sido imposta uma limitação à **data** de publicação, a presença de modelos mais estruturados para a aplicação da mineração de processos concentra-se nos últimos anos: 2018 (23%), 2017 (18%) e 2016 (14%). Isso sugere uma crescente na estruturação metodológica da aplicação de projetos de mineração de processos, corroborando a validade da presente pesquisa.

Dentre os **tipos de aplicação**, a área da saúde apresenta-se como a área mais representativa (32%) com relação ao volume de artigos de aplicação. Uma possível causa para essa presença maior pode dever-se à cultura de padronização de

processos presente na saúde, o que acaba refletindo nos projetos desenvolvidos na área.

No que diz respeito ao modelo propriamente dito, o mais frequente (a partir das FIGURA 10 a FIGURA 30) é o que foi desenvolvido por van Eck (2015) ou alguma versão simplificada dele. Com relação a quais **etapas do ciclo PDCA** estão contempladas na metodologia, os resultados mostram que apenas 50% das metodologias consideram o processo de planejamento (P). Essa baixa presença do planejamento é um ponto negativo no projeto como um todo, uma vez que o planejamento mal executado não permite a execução apropriada e o controle de processos ou objetivos do projeto (GLOBERSON; ZWIKAEEL, 2002).

No caso do desenvolvimento (D), existe uma aderência de 100% pelas metodologias. Possivelmente devido ao fato de ser uma metodologia de aplicação de algo, logo, é necessário que o desenvolvimento esteja presente no processo. O processo de conferência (C) está contemplado em 86% das metodologias apresentadas, ilustrando uma possível preocupação do alinhamento do que foi produzido para com o planejado. Por fim, a ação (A) aparece em apenas 45% dos casos, sugerindo que, após o desenvolvimento (D) e conferência (C), não são tomadas ações para que o quadro identificado seja corrigido.

Com relação à dimensão de análise da **ênfase**, 95% dos modelos apresentaram o processo como foco. As metodologias voltadas para o algoritmo figuram com 41% e as para pessoas contabilizam 45%.

As metodologias que possuem foco no *algoritmo*, conforme a leitura dos manuscritos, transmitiram uma presença secundária das outras duas ênfases; foram excessivamente voltadas para o algoritmo e a menção aos processos e/ou às pessoas foi vaga. No caso dos *processos*, é compreensível que praticamente todos as metodologias apresentem foco no processo, uma vez que são aplicações em casos reais. Por outro lado, a presença de apenas 45% do foco nas *pessoas* traz preocupação com relação à aderência dos desenhos de processo com o processo efetivamente desenvolvido. A gestão do conhecimento tácito e implícito é útil para o processo de desenho de processos, uma vez que captura elementos das etapas operacionais do processo (HERDER; WEIJNEN, 1999).

Por fim, a classificação com relação à iteratividade da metodologia diz respeito à possibilidade de correção/refinamento sem que seja necessário iniciar uma nova instância da metodologia. Foi avaliado se existia alguma etapa iterativa do processo

completo ou apenas em alguma etapa chave. Foi então constatado que apenas 27% das metodologias são iterativas. Esse valor, relativamente baixo, permite a interpretação de que a execução de apenas uma instância da metodologia seria o suficiente para que o problema fosse solucionado ou que o processo deveria ser repetido, mas como um novo projeto. O que não é uma interpretação necessariamente correta.



## 5 DESENVOLVIMENTO

Para o desenvolvimento, as características definidas como sendo ou não desejáveis para uma metodologia que seja compatível com o contexto da saúde partiram das dimensões presentes no QUADRO 5. Dessa forma, essa etapa consistiu em identificar a metodologia/combinção de metodologias, dentre as levantadas, que pudesse atender às características levantadas (Área de aplicação, Etapas do ciclo PDCA, Ênfase, “É iterativo?”). Além disso, as informações e comentários levantados pelo QUADRO 3 sobre elementos importantes em um projeto de mineração de processos foram incorporados.

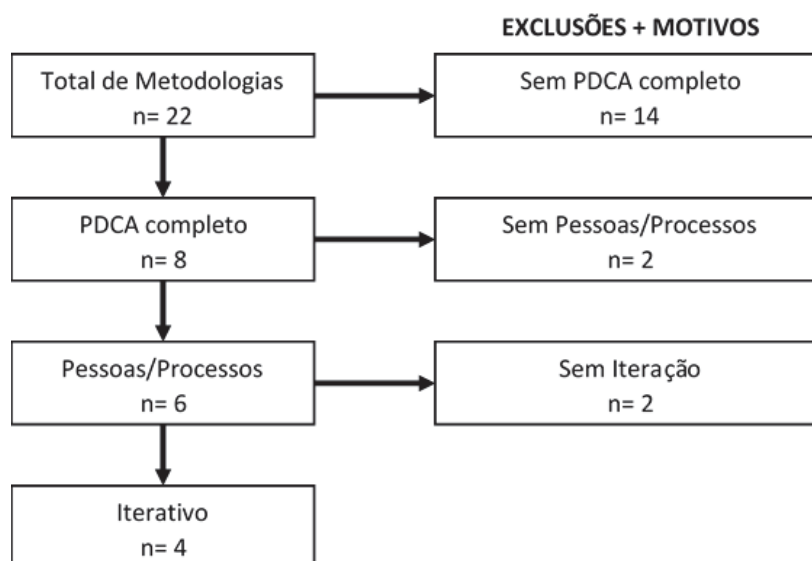
### 5.1 SELEÇÃO DA METODOLOGIA

Partindo das metodologias levantadas de mineração de processos e também das características definidas para sua análise, foi possível realizar a seleção da metodologia base a ser utilizada para proposição de uma para aplicação na área da saúde. Desta forma, para ser considerada uma metodologia “adequada” e ser utilizada como base, ela deveria contemplar alguns aspectos.

Inicialmente foram selecionadas aquelas que passam por todas as etapas do ciclo PDCA. Essa opção foi feita para que haja um processo completo; do planejamento à ação. O segundo critério de seleção foi a ênfase ser associada a *processos e pessoas*. A ênfase em algoritmos foi excluída pelo fato já comentado de que, uma vez que a sequência de passos da metodologia foi definida, qualquer algoritmo pode encaixar-se sem prejuízo. Por fim, foram selecionadas as metodologias que apresentam iteração em algum nível. Os resultados da filtragem ao longo de cada etapa estão apresentados na FIGURA 32.

Após o processo de filtragem apresentado na FIGURA 32, foram identificadas quatro metodologias que se enquadraram nos critérios. São elas as apresentadas por: WEERDT et al.(2013), VAN ECK et al. (2015), EPURE et al. (2016), ROJAS et al. (2016). A partir dessa seleção, cada uma foi explorada com o prisma das características do ambiente de saúde, sendo selecionada a apresentada por VAN ECK (2015) (*Process Mining Process Methodology* (PM<sup>2</sup>)). As outras três metodologias excluídas estão apresentadas (traduzidas) com o motivo da não adequação descrito.

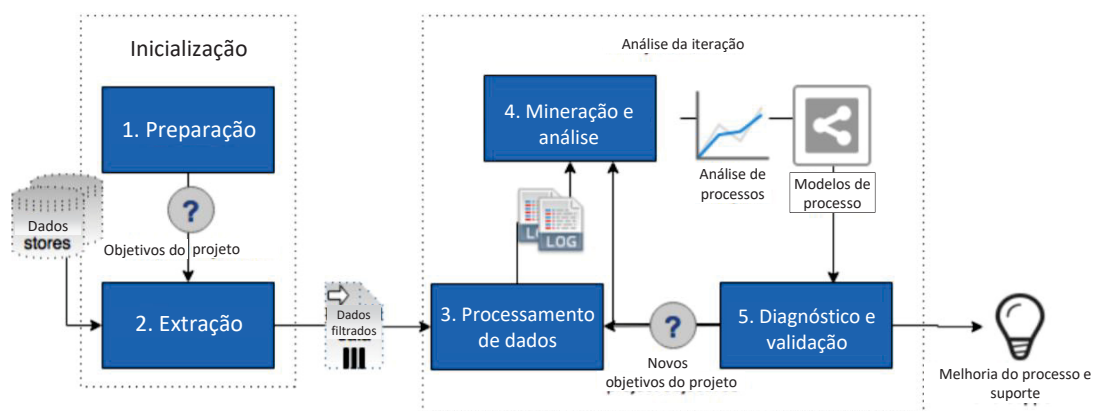
FIGURA 32 - FILTRAGEM DAS METODOLOGIAS



FONTE: O Autor (2018)

A metodologia apresentada por EPURE et al. (2016) encontra-se apresentada na FIGURA 33. O motivo pelo qual ela não foi selecionada, é o fato de ser uma versão simplificada da metodologia PM<sup>2</sup> proposta por VAN ECK et al. (2015). Dessa forma, foi feita a opção pelo modelo original.

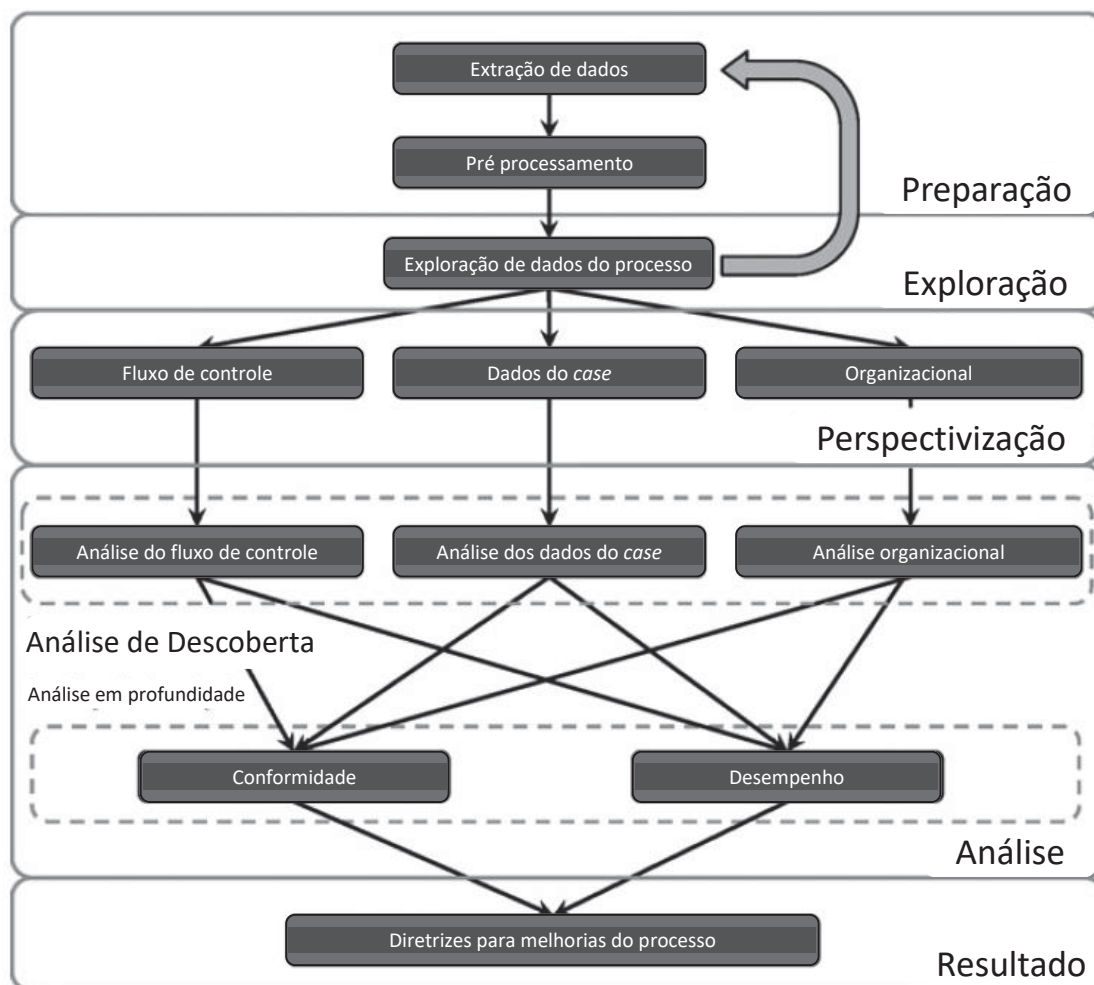
FIGURA 33 - METODOLOGIA DE EPURE et al. (2016) TRADUZIDA



FONTE: EPURE et al. (2016), tradução nossa

Em seguida, a metodologia apresentada em WEERDT et al. (2013) é apresentada na FIGURA 34. Apesar de mostrar-se como uma metodologia iterativa, tal fato só ocorre nas etapas iniciais do processo. Além disso, a descrição das etapas iterativas não detalha o que deve ser feito de forma objetiva, dificultando a reprodutibilidade e padronização.

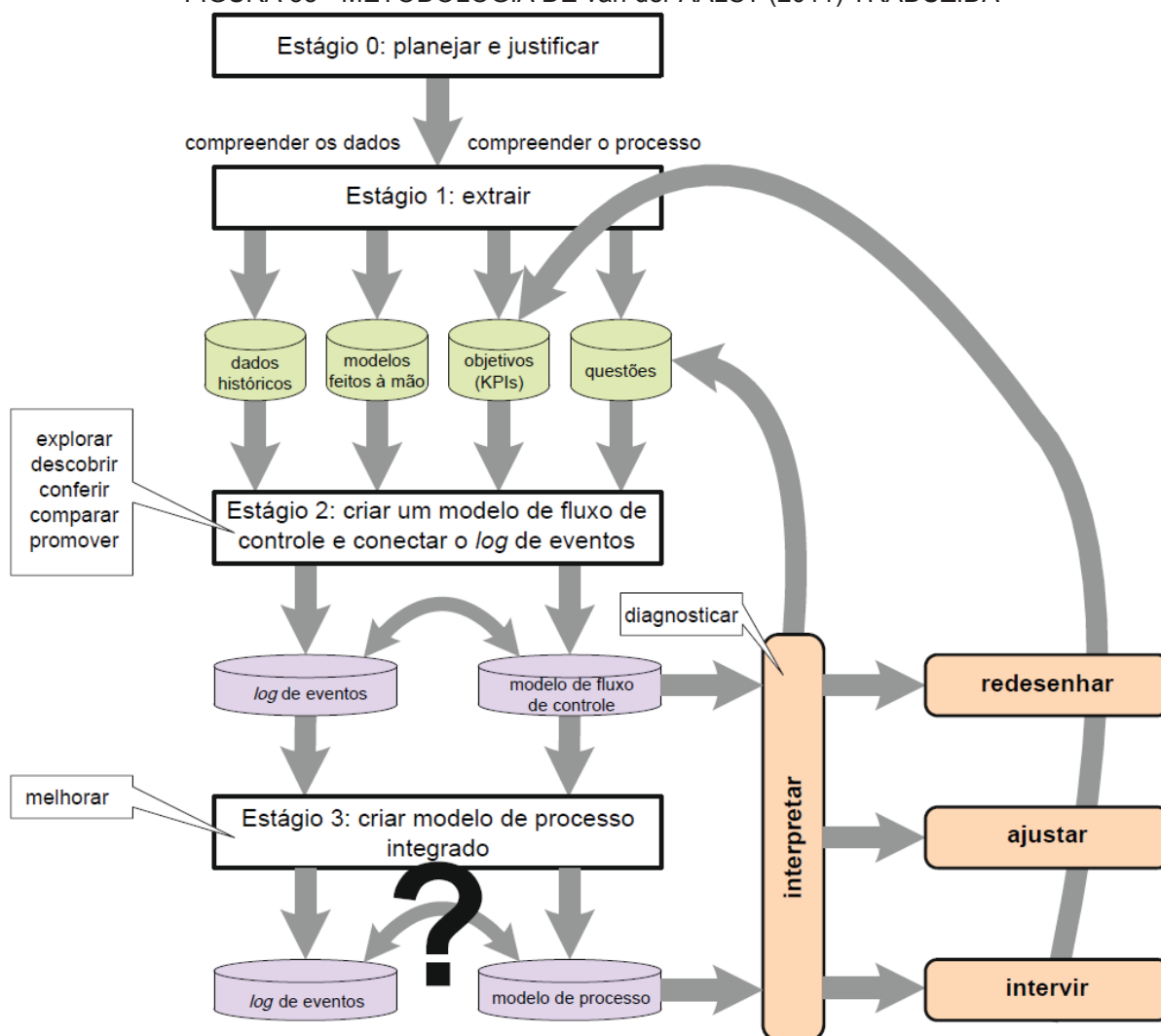
FIGURA 34 - METODOLOGIA DE WEERDT et al. (2013) TRADUZIDA



FONTE: WEERDT et al. (2013), tradução nossa

Por fim, a metodologia *L\* lifecycle model* apresentada em ROJAS et al. (2016), de autoria de VAN DER AALST (2013a) encontra-se apresentada na FIGURA 35. Essa metodologia seria uma opção potencial, uma vez que lida bem com todos os critérios elencados. Porém, conforme descrito em VAN DER AALST (2013a) e VAN ECK et al. (2015), não lida muito bem com processos do tipo “espaguete”, que é uma característica da área da saúde.

FIGURA 35 - METODOLOGIA DE van der AALST (2011) TRADUZIDA



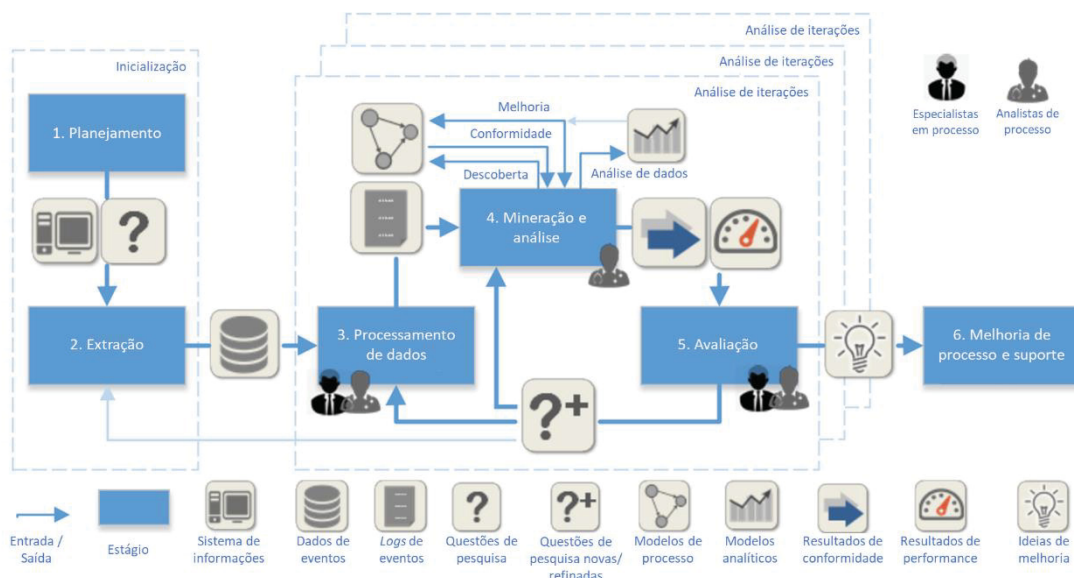
FONTE: (ROJAS et al., 2016 tradução nossa)

Com a seleção da metodologia, é apresentada a descrição de cada uma de suas etapas para que, na próxima seção, ela seja adaptada para a aplicação no ambiente de saúde.

#### 5.1.1 Process Mining Process Methodology (PM<sup>2</sup>)

A sequência de etapas desenvolvidas na metodologia PM<sup>2</sup> está apresentada na FIGURA 36. Em seguida, é apresentada a tradução livre (realizada pelo autor) da descrição de cada uma dessas etapas, conforme o artigo em que a PM<sup>2</sup> foi apresentada (VAN ECK et al., 2015).

FIGURA 36 - METODOLOGIA DE VAN ECK (2015) TRADUZIDA



FONTE: (VAN ECK et al., 2015 tradução nossa)

## 1. Planejamento

O objetivo da etapa de planejamento é configurar o projeto e determinar as questões de pesquisa. Consideram-se duas metas principais para iniciar projetos de mineração de processos: melhoria do desempenho de um processo ou verificação de conformidade com determinadas regras e regulamentações.

As entradas dessa etapa são os processos da organização. As saídas são as questões de pesquisa relacionadas ao objetivo e um conjunto de sistemas de informação que apoiem a execução dos processos a serem analisados.

Foram identificadas três atividades para esta etapa: seleção de processos, identificação das questões de pesquisa (Q.P.) e composição de equipe de projeto. A ordem em que essas atividades são executadas pode variar, uma vez que pode haver um objetivo específico ou uma questão de pesquisa presente antes do início do projeto de mineração de processos.

- a. **Seleção de processos**: Um projeto de mineração de processo geralmente começa com a seleção dos processos de negócios a serem analisados e aprimorados. Tanto as características do processo quanto a qualidade dos dados dos eventos devem ser levadas em consideração, uma vez que têm grande influência nos resultados finais do projeto (BOSE; MANS; VAN DER AALST, 2013). Além desses dois fatores, também se considera a mutabilidade dos processos de uma organização, ou seja, a organização precisa ser capaz de influenciar ou adaptar

as execuções de processos com base nos resultados. Isso é importante se a melhoria do processo for o objetivo principal do projeto. Depois de selecionar os processos, os sistemas de informação que armazenam os dados relevantes do processo são identificados.

- b. Identificação das questões de pesquisa: Durante esta atividade, os objetivos são identificados e traduzidos em questões de pesquisa, que são definidas como questões relacionadas ao processo selecionado e que podem ser respondidas usando dados de eventos. As questões de pesquisa podem estar relacionadas a diferentes aspectos dos processos de negócios como, por exemplo: qualidade, tempo, recursos, custo. Diversos estudos de caso (SURIADI et al., 2013) demonstraram a importância de definir questões de pesquisa concretas para um projeto de mineração de processos bem-sucedido. Porém, foi demonstrado, no estudo de caso presente no artigo de apresentação do PM<sup>2</sup>, que questões de pesquisa abstratas da fase de inicialização podem ser refinadas através da análise exploratória, resultando em ideias concretas de melhoria e *insights* valiosos.
- c. Composição da equipe de projeto: A última atividade envolve selecionar as pessoas que irão trabalhar no projeto. Estudos de caso anteriores (incluindo o estudo de caso do artigo) mostram que as equipes de projeto precisam de especialistas com diferentes históricos (SURIADI et al., 2013). Foram definidas as seguintes funções: gestores do negócio (responsáveis pelos processos de negócios), especialistas em negócios (que conhecem o negócio e as execuções dos seus processos), especialistas em sistemas (que estão familiarizados com o aspecto de TI<sup>21</sup> dos processos e dos sistemas que apoiam os processos) e analistas de processos (que são especializados em analisar processos e aplicar técnicas de mineração de processos). As funções mais importantes são os especialistas em negócios e os analistas de processos, cuja colaboração é essencial para avaliar as descobertas da análise e garantir que sejam relevantes e aplicáveis.

## 2. Extração

A etapa de extração tem como objetivo extrair dados de eventos e, opcionalmente, modelos de processos. As entradas para este estágio são as questões

---

<sup>21</sup> Tecnologia da Informação

de pesquisa e os sistemas de informação que suportam a execução dos processos selecionados para serem analisados. As saídas desse estágio são dados de eventos, ou seja, uma coleção de eventos sem noção predefinida de caso ou classes de eventos e, possivelmente, modelos de processo.

Foram identificadas três atividades para esta etapa: determinação do escopo, extração de dados de eventos e transferência de conhecimento do processo.

- a. Determinação do escopo: Essa atividade envolve a determinação do escopo da extração de dados, com a qual os dados do evento serão criados. São dados quatro exemplos de questões a serem consideradas: (1) com qual granularidade devem ser extraídos os dados do evento (por exemplo, considerando eventos relacionados a pedidos, mas negligenciando eventos relacionados aos itens dos pedidos); (2) dentro de qual período; (3) quais atributos de dados devem ser extraídos; (4) qual correlação entre os dados deve ser usada para coletá-los.
- b. Extração de dados de eventos: Depois que o escopo de extração é determinado, os dados de eventos podem ser criados por meio da coleta dos dados relacionados aos processos selecionados. Tais dados são coletados dos sistemas de informações relevantes e consolidados em uma única coleção de eventos, por exemplo, uma tabela na qual cada entrada representa um evento.
- c. Transferência de conhecimento do processo: Essa atividade pode ser executada simultaneamente com a criação de dados do evento. O conhecimento tácito relacionado aos processos de negócios selecionados e aos atributos de dados é trocado entre especialistas de negócios e analistas de processo, por meio de entrevistas ou sessões de *brainstorming*, permitindo que os analistas sejam mais efetivos nas etapas de processamento de dados e mineração. Tal conhecimento de processo pode incluir documentação de processo escrita ou modelos de processo feitos à mão. O conhecimento do processo é compartilhado em todo o projeto, mas a compreensão do processo é essencial para que o estágio de processamento de dados seja eficaz.

Em contraste com as metodologias de mineração de processos existentes, a extração de dados do evento foi dividida da criação e processamento do *log*. Uma razão é que a extração de dados de eventos é demorada e menos frequentemente repetida do que as atividades de processamento de dados, como a filtragem (NOOIJEN; DONGEN; FAHLAND, 2013). Outra razão é que é possível criar diferentes



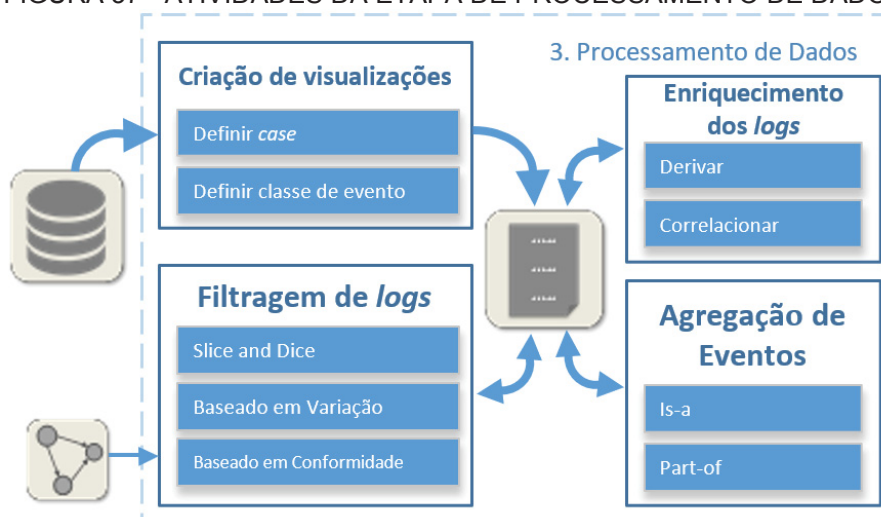
visualizações dos mesmos dados de eventos que resultam em diferentes *logs* de eventos, como será discutido na próxima seção.

### 3. Processamento de dados

O principal objetivo do estágio de processamento de dados é criar *logs* de eventos como diferentes visualizações dos dados de eventos obtidos e processá-los de forma que sejam ideais para o estágio de mineração e análise. Além dos dados do evento como nossa entrada principal, também é possível usar modelos de processo como uma entrada para filtrar os dados do evento. As saídas são *logs* de eventos que são usados no estágio de mineração e análise.

Quatro tipos de atividades estão identificados neste estágio: criação de visualizações, agregação de eventos, enriquecimento de *logs* e filtragem de *logs*. A FIGURA 37 mostra uma visão geral dessas atividades e como elas são aplicadas.

FIGURA 37 - ATIVIDADES DA ETAPA DE PROCESSAMENTO DE DADOS



FONTE: (VAN ECK et al., 2015 tradução nossa)

- a. **Criação de visualizações:** Os *logs* de eventos são exibições específicas nos dados do evento, criados pela definição de noções de *cases* e classes de eventos. A noção de *case* relaciona eventos de tal forma que, juntos, eles formam uma instância de processo, enquanto que classes de eventos distinguem atividades diferentes dentro de uma instância de processo. A visão criada depende do objetivo da análise, por exemplo uma ordem é uma noção de *case* lógica para analisar os tempos de atravessamento, enquanto o recurso é uma noção de *case* melhor para analisar a utilização de recursos. Uma situação semelhante é válida para definir as classes de evento: se cada ordem for uma instância de processo e



o recurso envolvido em um evento for sua classe de evento, os algoritmos de descoberta de processo produzirão gráficos de transferência de trabalho.

- b. Agregação de eventos: A agregação de eventos pode ajudar a reduzir a complexidade e melhorar a estrutura dos resultados de mineração (BOSE; VAN DER AALST, 2009). A agregação foi dividida em dois tipos: “*is-a*” e “*part-of*”. A agregação *is-a* considera diferentes tipos de eventos pertencentes a uma classe de eventos equivalente, mas mais geral, enquanto o número de eventos permanece o mesmo. Por exemplo, dois eventos rotulados com “Análise Manual Simples” e “Análise Manual Complexa” são considerados instâncias da classe de evento “Análise Manual”, porém, permanecem como dois eventos. Por outro lado, a agregação *part-of* mescla vários eventos em eventos maiores, como é o caso dos subprocessos. Ambos os tipos de agregação também podem ser aplicados em sentido inverso (definindo uma especialização). Uma técnica mais geral é definir uma hierarquia baseada em atributos de eventos que podem ser usados para agregar eventos, como é discutido em (VAN DER AALST, 2013b) (ex: considerando a localização em um nível de cidade, país ou continente)
- c. Enriquecimento dos logs: Os *logs* de eventos, como quaisquer outros dados, podem ser enriquecidos com vários atributos adicionais (LEONI; VAN DER AALST; DEES, 2014). São apresentadas duas maneiras de enriquecer um *log* de eventos: (1) derivar ou computar eventos e atributos de dados adicionais com base no próprio *log* ou (2) adicionar dados externos (correlacionar). O tempo de processamento de um *case* pode ser um atributo de dados calculado, enquanto que a adição de informações sobre o clima no momento em que um evento ocorreu é um exemplo de inclusão de dados externos.
- d. Filtragem de logs: Por fim, a filtragem é uma etapa de processamento de dados bem conhecida e usada com frequência para reduzir a complexidade ou focar a análise em uma parte específica do conjunto de dados. Essa atividade geralmente é executada várias vezes em uma iteração de análise para obter diferentes perspectivas sobre os dados do evento. São elencadas três técnicas de filtragem: *slice and dice* (também conhecida como filtragem de atributos), baseada em variação e baseada em conformidade.
  - i. *Slice and dice* - pode ser usada para remover eventos ou traços com base nos valores registrados para um atributo específico (ex: nome da atividade, identificador de recurso ou *timestamps* de eventos, ou com base em

estatísticas simples - número de eventos de uma duração de rastreo ou caso).

- ii. *Filtragem baseada em variação* - agrupa traços semelhantes (ex: *clustering*) que podem ser usados para particionar o *log* de eventos, a fim de descobrir modelos de processos mais simples para cada uma das partições de um processo complexo (REBUGE; FERREIRA, 2012).
- iii. *Filtragem baseada em conformidade* - pode ser usada para remover traços ou eventos que não se comprem uma determinada regra ou se não se ajustam a um determinado modelo de processo, sendo uma forma muito flexível de filtragem.

#### 4. Mineração e Análise

Na etapa de mineração e análise, aplicam-se técnicas de mineração de processo em *logs* de eventos e busca-se responder as perguntas de pesquisa e obter *insights* sobre o desempenho e a conformidade dos processos. Se as questões de pesquisa forem mais abstratas, podem ser combinadas técnicas exploratórias com as de descoberta de processos nos *logs* de eventos para obter uma visão geral do processo de negócios (ex: fluxo de controle). Uma vez que questões específicas de pesquisa tenham sido definidas, a análise pode se concentrar em responder a perguntas concretas de pesquisa (ex: “a diferença entre os tempos de atravessamento dos casos que executaram as atividades do tipo A e os casos que ignoraram essa atividade”).

As entradas para esta etapa são os *logs* de eventos. Além disso, se modelos de processo estiverem disponíveis, eles também podem ser usados para verificação de conformidade e atividades de aprimoramento. A saída para esta etapa são elementos que respondem às perguntas de pesquisa relacionadas aos objetivos de desempenho e conformidade.

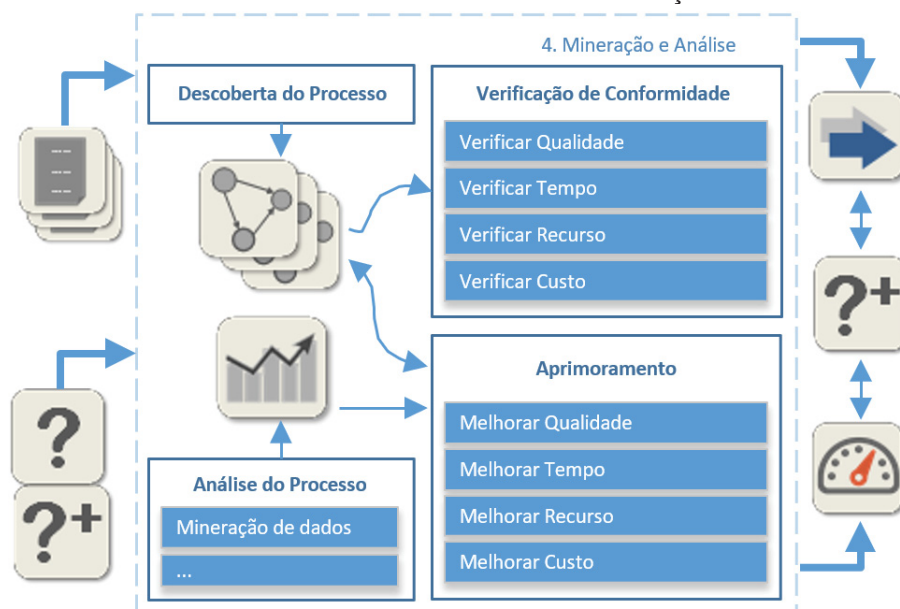
São identificadas quatro tipos de atividades para esta etapa: descoberta de processos, verificação de conformidade, aprimoramento e análise de processos<sup>22</sup>. As três primeiras atividades são técnicas de mineração de processos bem conhecidas (VAN DER AALST, 2011). A análise do processo é outra técnica de análise

---

<sup>22</sup> *Process Analytics*

complementar (ex: mineração de dados e análise visual, que podem ser aplicados no contexto de processos de negócios) (LEONI; VAN DER AALST; DEES, 2014). A FIGURA 38 mostra uma visão geral das quatro atividades.

FIGURA 38 - ATIVIDADES DA ETAPA DE MINERAÇÃO E ANÁLISE



FONTE: (VAN ECK et al., 2015 tradução nossa)

- a. Descoberta de Processos: Dado um *log* de eventos como entrada, geralmente inicia-se com a aplicação de técnicas de descoberta de processo, que retornam um modelo de processo baseado em fatos como saída<sup>23</sup>.
- b. Verificação de conformidade: Dado um modelo de processo - descoberto ou documentado - descrevendo o comportamento desejado e um *log* de eventos que registra o comportamento real, as técnicas de verificação de conformidade visam a detecção de inconsistências entre um modelo de processo e seu *log* de execução correspondente (ROZINAT; VAN DER AALST, 2008). As questões de pesquisa relacionadas à conformidade dos processos de negócios a diferentes aspectos (qualidade, tempo, recurso, custo, etc.) podem ser verificadas usando técnicas de verificação de conformidade, cujos resultados também podem ser usados para aprimorar os modelos de processo.
- c. Análise do processo: Além das três atividades de mineração de processo, outras técnicas de análise podem ser aplicadas no contexto de *logs* de eventos e modelos de processos, como técnicas de mineração de dados (LEONI; VAN DER AALST; DEES, 2014) ou análises visuais (ex: histogramas de eventos por caso), cujos

<sup>23</sup> Para discussões sobre diferentes técnicas de descoberta de processos, ver, e. [7]

resultados podem ser usados para aprimorar modelos de processo com aspectos adicionais.

- d. **Aprimoramento:** A atividade de aprimoramento é definida como a extensão ou melhoria de um modelo de processo existente usando informações sobre o processo real registrado em um *log* de eventos (VAN DER AALST, 2011). Pode ocorrer, por exemplo, estendendo o modelo de processo com informações de desempenho relacionadas a tempo ou custo ou reparando o modelo de processo às execuções atuais mostradas pelo *log* de eventos correspondente. Os resultados do aprimoramento são modelos de processo (ex: aprimorado com diferentes aspectos) ao passo que os resultados da verificação de conformidade podem ser considerados sem qualquer modelo de processo.

## 5. Avaliação

O objetivo da etapa de avaliação é relacionar as conclusões da análise com as ideias de melhoria que vão ao encontro dos objetivos do projeto. As entradas são os modelos de processo, achados relacionados a desempenho e conformidade do estágio de análise. As saídas são ideias de melhoria ou novas questões de pesquisa.

As atividades para este estágio são: Diagnosticar e Verificar e Validar (V & V).

- a. **Diagnosticar:** Diagnosticar os achados obtidos através da mineração e análise inclui: (1) interpretar corretamente os resultados (ex: compreender o modelo de processo descoberto); (2) distinguir resultados interessantes ou incomuns dos esperados (ex: grande conjunto de execuções anormais); (3) identificar ou refinar questões de pesquisa para possíveis novas iterações.
- b. **Verificar e validar (V&V):** A exatidão das descobertas (inesperadas) é investigada. A verificação compara os achados obtidos com os dados originais e as implementações do sistema, enquanto que a validação compara os achados com as reivindicações das partes interessadas no processo (ex: entrevistando os recursos envolvidos nos processos). Tanto a verificação quanto a validação podem ajudar a identificar as causas-raiz subjacentes e projetar ideias para possíveis melhorias no processo.

Um dos desafios em projetos de mineração de processo é que os analistas de processo não são especialistas no processo em que estão analisando (BOZKAYA; GABRIELS; WERF, 2009; SURIADI et al., 2013), o que significa que eles podem ter

dificuldades em determinar as causas dos resultados inesperados da análise. Portanto, é essencial que especialistas em processos estejam envolvidos na verificação e validação dos resultados. O ideal é que eles já estivessem envolvidos durante o estágio de mineração anterior, guiando a análise para garantir que os resultados sejam úteis para a organização.

## 6. Melhoria de Processos e Suporte

O objetivo do estágio de melhoria e suporte de processos é usar os *insights* obtidos para modificar a execução real do processo. As entradas deste estágio são as idéias de melhoria do estágio de avaliação. As saídas deste estágio são modificações no processo.

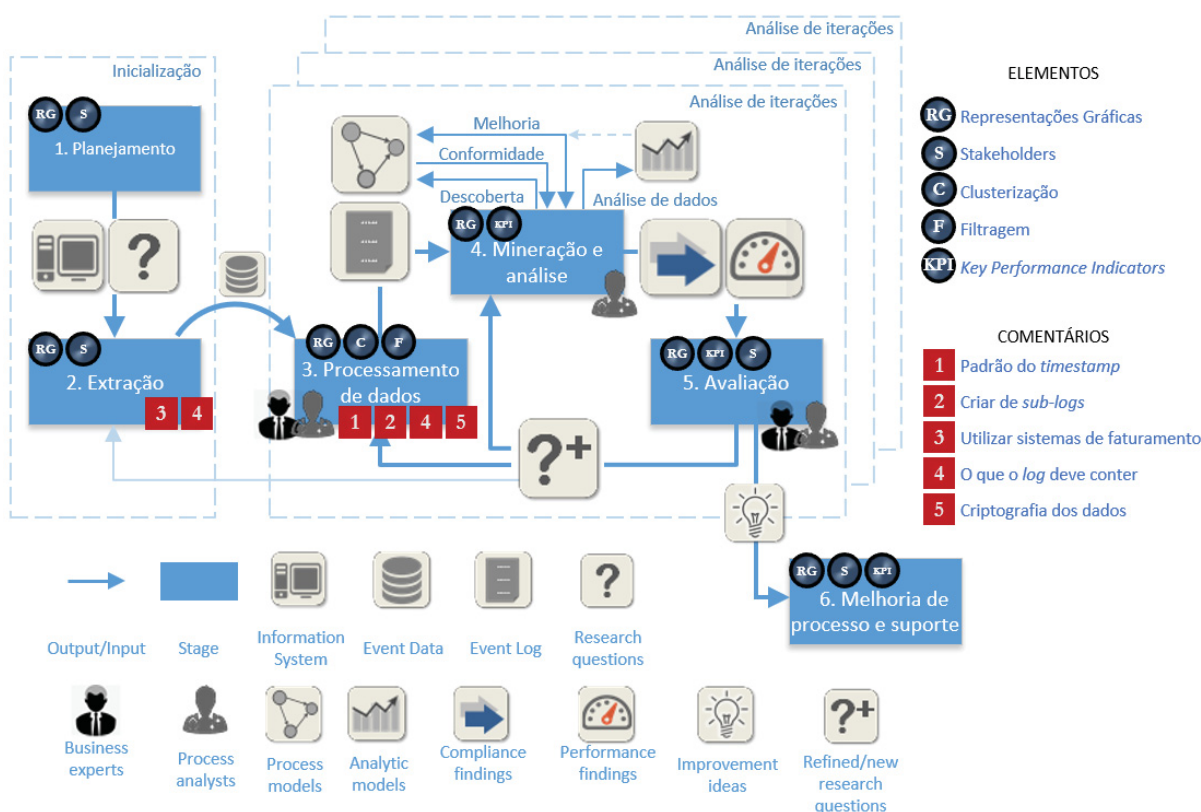
As atividades são: implementação de melhorias e suporte às operações.

- a. Implementando melhorias: A obtenção de melhorias no processo é geralmente a principal motivação para um projeto de mineração de processos. No entanto, a implementação efetiva das modificações do processo é geralmente um projeto separado com uma área de especialização diferente. Os resultados de um projeto de mineração de processos formam a entrada - baseada em fatos - dos esforços de melhoria de processo. Abordagens que enfocam esta área incluem reengenharia de processos e Six Sigma (HARMON, 2007). Uma vez que houve a alteração do processo, as melhorias podem ser medidas em um projeto de análise à parte.
- b. Suporte de operações: A mineração de processos pode fornecer suporte operacional por meio da detecção de casos com execução problemática, prevendo seu futuro ou sugerindo ações recomendadas. Para usar a mineração de processo para suporte operacional, é essencial que os resultados sejam de alta qualidade e que exista uma infraestrutura de TI que vincule esses resultados a dados de eventos ativos. Tal vinculação é uma forma desafiadora de mineração de processos, adequada apenas para processos muito estruturados (VAN DER AALST, 2011).

## 5.2 APRESENTAÇÃO DA METODOLOGIA PM<sup>2</sup>HC

Com as etapas da metodologia de VAN ECK et al. (2015) descritas, as dimensões analisadas previamente na seção 4.2.1 *Estado da arte e aplicações da mineração de processos*, foram adicionadas na etapa de uso sugerido; Representações Gráficas (RG), *Stakeholders* (S), Filtragem (F), Clusterização (C) e Indicadores de Performance (KPI). Além disso, as observações pontuais extraídas na mesma seção estão dispostas na etapa de uso; (1) padrão do *timestamp*, (2) criação de sub-logs, (3) utilização de dados do sistema de faturamento, (4) o que um *log* deve conter, (5) criptografia dos dados. Essa metodologia foi chamada de *Process Mining Process Methodology in Healthcare* (PM<sup>2</sup>HC), para indicar a metodologia (PM<sup>2</sup>) e a área de atuação (*Healthcare*). A FIGURA 39 apresenta as etapas de inserção dos complementos.

FIGURA 39 – PROCESS MINING PROCESS METHODOLOGY IN HEALTHCARE



FONTE: (Adaptado de VAN ECK et al., 2015).

A escolha dos momentos do projeto em que deveriam estar cada um dos elementos identificados ocorreu por meio da análise das características presentes na descrição de cada etapa do PM<sup>2</sup>. Cada etapa foi confrontada com suas necessidades



e carências que poderiam vir a ser solucionadas com o acréscimo do que foi identificado na RSL.

Essas inserções, que ocorreram em todas as etapas do PM<sup>2</sup>, buscam deixar a metodologia melhor adaptada para o ambiente de saúde, uma vez que é uma área com demandas muito específicas (tanto com relação ao processo em si, quanto com relação ao seu papel social). Dessa forma, o PM<sup>2</sup>HC busca ser um norte para os pesquisadores no desenvolvimento de projetos de mineração de processos em saúde. A seguir, serão justificadas as inserções em cada etapa.

A etapa de PLANEJAMENTO lida com a configuração do projeto e com a determinação das questões de pesquisa. Para que o projeto consiga desenvolver-se de forma a alinhar-se com as expectativas da esfera gerencial e dos financiadores, o elemento *Stakeholders* foi associado à essa etapa justamente por haver a necessidade de estar atento às suas demandas. Junto a isso, a presença de Representações Gráficas se faz necessária pois permite ilustrar aspectos complexos de uma organização de uma maneira simples e poderosa ao mesmo tempo em que essa apresentação condensada e mais adequada à camada gerencial dos resultados amplia sua aceitação e usabilidade (MEYER, 1991; RIVERA; DURAN, 2004). Em suma, as representações gráficas buscam auxiliar a equipe do projeto, formada por pessoas com formações distintas, a chegar a um consenso sobre o processo e os objetivos envolvidos. A metodologia PM<sup>2</sup> não menciona tais elementos, o que deixa o planejamento muito abrangente e dificulta/posterga a definição de objetivos que vão ao encontro das necessidades da instituição.

Para a etapa de EXTRAÇÃO (que busca a extração dos dados de eventos do sistema e, opcionalmente, modelos de processos), a mesma necessidade da integração dos *Stakeholders* e do uso de Representações Gráficas se faz presente. O primeiro pela necessidade de determinação do escopo e o segundo pela facilitação da transferência do conhecimento. Além disso, a opção pela extração dos dados do sistema de faturamento/cobrança é algo a ser considerado, dada a maior qualidade dos dados (RAMOS, 2009). Da mesma forma, um *log* deve conter ID do paciente, nome da atividade, departamento envolvido, *timestamp* e, opcionalmente, atributos para desenvolver análises adicionais (RAMOS, 2009). A metodologia PM<sup>2</sup> não estabelece a necessidade de alinhamento com os *Stakeholders* e não utiliza/sugere nenhum tipo de ferramenta de auxílio nesse alinhamento. Além disso, não contempla sugestões relacionadas ao formato dos *logs*, nem do sistema de extração.

Na etapa de PROCESSAMENTO DE DADOS, ocorre a criação de *logs* de eventos como diferentes visualizações dos dados de eventos obtidos e processá-los de forma que sejam ideais para o estágio de mineração e análise. Nessa etapa, os *Stakeholders* já não estão mais envolvidos, por ser mais técnica (envolvendo analistas de TI, processos e de mineração). Nessa etapa, bem como nas subsequentes, as Representações Gráficas são utilizadas como facilitador no entendimento do processo. Dois elementos levantados na revisão bibliográfica que já se encontram contemplado nessa etapa são a Filtragem e a Clusterização (chamado de “agregação de eventos”). O que ocorre nesses casos, é a ênfase na sua importância; uma vez que diversos artigos os citaram.

Os comentários adicionais associados à essa etapa dizem respeito à diversos pontos de sugestão durante o processamento dos dados. Inicialmente, o *timestamp* no formato “dd-MM-aaaa HH:mm” facilita o ordenamento das atividades. Para lidar com modelos do tipo “espaguete” recomenda-se que o *log* seja quebrado em dois ou mais *sub-logs* (para reduzir a complexidade). Os elementos recomendados de estarem contidos no *log* são: ID do paciente, nome da atividade, departamento envolvido, *timestamp* e, opcionalmente, atributos para desenvolver análises adicionais (como idade do paciente, por exemplo). Desse comentário, apenas a questão do atributo está mencionada no PM<sup>2</sup>, no enriquecimento dos *logs*. Por fim, o fato de os *logs* serem compostos por dados sensíveis (informações de pacientes, cobrança, etc.), sugere criptografá-los. A metodologia PM<sup>2</sup> não explora o uso de nenhum tipo de tratamento de dados (ex: filtragem e clusterização) e também não apresenta a necessidade de alinhamento do entendimento dos profissionais com diferentes cargos/formações. Além disso, apesar de ter sido desenvolvida como uma alternativa ao *L\* lifecycle model* para conseguir lidar com modelos do tipo “espaguete”, não apresenta alternativas sobre como lidar com esse tipo de modelo quando da realização do projeto.

A etapa de MINERAÇÃO E ANÁLISE busca responder as perguntas de pesquisa e obter *insights* sobre o desempenho e a conformidade dos processos. Nessa etapa, o uso das Representações Gráficas mantém-se importantes, especialmente para o momento de análise de conformidade e análise do processo. Além disso, o uso de KPIs (indicadores de desempenho) fornece elementos concretos para que a análise seja realizada de maneira precisa, em especial nas etapas de verificação de conformidade, aprimoramento e análise do processo. De modo geral,



os KPIs permitem avaliar se os resultados encontram-se alinhados com a estratégia organizacional e é usado pelos gestores para analisar o andamento do projeto e diagnosticar a organização (SEIFY, 2010). O PM<sup>2</sup> menciona brevemente o uso de Representações Gráficas (como “análise visual”) apenas na etapa de *Análise do Processo*. Já no caso dos KPIs, existem as etapas de “verificação de conformidade” e “aprimoramento”, que sugerem indiretamente o uso de métricas para análise. Dessa forma, foi considerado procedente incluir os KPIs nessa etapa para explicitar a validade de seu uso.

A etapa de AVALIAÇÃO busca relacionar as conclusões da análise com as ideias de melhoria que vão ao encontro dos objetivos do projeto. Nessa etapa, ocorre o retorno da presença do *Stakeholder* para poder avaliar se as conclusões e ideias de melhoria estão efetivamente alinhadas com os objetivos do projeto. Esse alinhamento acaba utilizando, por exemplo, os KPIs gerados nas etapas anteriores, ao mesmo tempo em que possui suporte das Representações Gráficas para permitir uma melhor interpretação. Nessa etapa, o PM<sup>2</sup> não menciona o uso de métricas desenvolvidas ao longo do projeto, tampouco o envolvimento da camada gerencial para validação do cumprimento dos objetivos. A questão das Representações Gráficas também é deixada de lado nessa etapa.

Por fim, na etapa de MELHORIA DE PROCESSO E SUPORTE foram mantidos os mesmos elementos presentes na etapa de AVALIAÇÃO. A presença do *Stakeholder* desempenha o papel do elemento norteador, buscando o cumprimento do objetivo. As Representações Gráficas são utilizadas para nivelar o entendimento da equipe acerca do projeto com o uso dos KPIs gerados nas etapas anteriores e que se encontram adequados à etapa de melhoria. Aqui há a ressalva de que essa interpretação está sendo feita para a etapa de MELHORIA DE PROCESSOS E SUPORTE contida dentro do mesmo projeto de mineração de processos. No caso de haver um projeto separado para a implementação das melhorias, o uso dos três elementos mencionados (*Stakeholders*, Representações Gráficas e KPIs) pode ser feito de forma isolada, adaptado ao contexto. O PM<sup>2</sup> não menciona nem o uso desses elementos citados e, tampouco, trata de alternativas de como lidar com esses tópicos.

Após passar pelas etapas do PM<sup>2</sup>, com as respectivas inserções dos elementos e comentários, foi possível compreender a validade do PM<sup>2</sup>HC, que foi explorado no estudo de caso apresentado na seção seguinte.

## 6 APLICAÇÃO

Para verificar a aderência do PM<sup>2</sup>HC a um projeto real de mineração de processo em saúde, foi feita sua aplicação na Unidade do Laboratório de Análises Clínicas (ULAC) do Complexo Hospital de Clínicas da Universidade Federal do Paraná (CHC/UFPR). Tal escolha foi feita devido à facilidade de acesso (tanto física quanto burocrática) e pela sua representatividade no contexto nacional.

### 6.1 O COMPLEXO HOSPITAL DE CLÍNICAS DA UNIVERSIDADE FEDERAL DO PARANÁ

O Hospital de Clínicas foi inaugurado em 5 de agosto de 1961 e, com quase 60 anos de história, pode orgulhar-se de ser o maior hospital público do Paraná e o terceiro hospital universitário federal do país (PROPLAN, 2017). Está inserido na 2ª Regional de Saúde (2ª RS) do Estado do Paraná, que engloba Curitiba e outros 28 municípios. Essa região metropolitana contabiliza 3.223.836 habitantes (oitava mais populosa do Brasil) e 16.581,21km<sup>2</sup> (segunda maior região metropolitana em extensão) (SECRETARIA DE DESENVOLVIMENTO URBANO, 2016).

Ao longo de sua trajetória, sempre se manteve uma posição de destaque pelas inovações e pelo perfil de atendimento (média e alta complexidade). Isso teve desdobramentos no registro e controle das atividades. No ano de 1979 houve a informatização dos serviços administrativos do hospital e, em 1989, é criada a Assessoria de Informática (atualmente Serviço de Informática e integrante da Unidade de Informação) que em conjunto com a IBM informatiza o hospital (EBSERH. EMPRESA BRASILEIRA DE SERVIÇOS HOSPITALARES, 2010, 2018).

Em 30 de outubro de 2014 houve a contratação da Empresa Brasileira de Serviços Hospitalares (EBSERH), que tem a finalidade de prestar apoio à universidade na gestão do hospital. Nesse momento, tanto o Hospital de Clínicas (HC), quanto a Maternidade Victor Ferreira do Amaral (MVFA), ambos hospitais universitários (HUs) pertencentes à Universidade Federal do Paraná, passaram a atuar sob a mesma esfera administrativa denominada de Complexo HC (CHC). O HC encontra-se inserido no Sistema Único de Saúde (SUS), com atendimentos de referência nos níveis terciário e quaternário. Também está articulado com a rede pública de saúde nos níveis primário e secundário (PROPLAN, 2017).

Sendo um hospital universitário, atende uma demanda de cerca de 3.760 alunos de graduação ao ano de diversos cursos da área da saúde: Medicina, Enfermagem, Nutrição, Farmácia, Odontologia, Terapia Ocupacional, Psicologia. Além disso, possui diversas atividades relacionadas ao ensino, pesquisa e extensão que se apresentam na forma de estágios supervisionados, atividades de extensão junto à comunidade assistida e pesquisa articulada com a assistência hospitalar (PROPLAN, 2017).

Além do que ocorre no âmbito da graduação, possui 51 programas de Residência Médica e 32 programas para Residência Multiprofissional na área da Saúde. Abriga também programas de Mestrado e Doutorado. Realiza quase três milhões de procedimentos por ano entre atendimentos ambulatoriais, cirurgias, internações e exames. Todo o atendimento é feito pelo SUS, sendo o maior prestador de serviços ao SUS no Paraná. São 545 leitos físicos em 49 unidades de internação, dos quais 400 estão ativos e contratados com o Sistema Único de Saúde (PROPLAN, 2017).

## 6.2 APLICAÇÃO

Para o desenvolvimento do estudo de caso, foram seguidos os passos do PM<sup>2</sup>HC, conforme apresentado na FIGURA 39. Uma apresentação das etapas, sub-etapas e tópicos a observar ao longo de cada etapa do PM<sup>2</sup>HC que foram seguidas no presente estudo de caso está disposto no QUADRO 6. A coluna “etapa” apresenta o projeto de uma maneira mais macro, sendo que um maior nível de granularidade se encontra nas sub-etapas da próxima coluna. Por fim, na coluna “Tópicos a observar”, estão dispostos os elementos identificados anteriormente na revisão da literatura (*Stakeholders*, Representações Gráficas, Clusterização, Filtragem e *Key Performance Indicators*) com as iniciais relacionadas a cada sub-etapa em que foram considerados como diferenciais na aplicação da mineração de processos na saúde. Ainda nessa coluna, os comentários (numerados de 1 a 5) encontram-se entre parênteses. Tanto no primeiro caso, quanto no segundo, estão dispostos em ordem decrescente de importância na sub-etapa.

QUADRO 6 - ETAPAS, SUB-ETAPAS E TÓPICOS A OBSERVARDO NO PM<sup>2</sup>HC

ETAPA	SUB-ETAPA	Tópicos a observar
6.2.1 Planejamento	6.2.1.1 Seleção de processos 6.2.1.2 Identificação das questões de pesquisa 6.2.1.3 Composição da equipe de projeto	S - RG S S
6.2.2 Extração	6.2.2.1 Determinação do escopo 6.2.2.2 Extração de dados de eventos 6.2.2.3 Transferência de conhecimento do processo	S - RG RG (4 – 3) RG
6.2.3 Processamento de dados	6.2.3.1 Criação de visualizações 6.2.3.2 Agregação de eventos 6.2.3.3 Enriquecimento dos logs 6.2.3.4 Filtragem de logs	RG (4 – 1 – 5) C – RG (2) RG F – RG (2)
6.2.4 Mineração e análise	6.2.4.1 Descoberta de Processos 6.2.4.2 Verificação de conformidade 6.2.4.3 Análise do processo 6.2.4.4 Aprimoramento	RG KPI - RG KPI - RG KPI - RG
6.2.5 Avaliação	6.2.5.1 Diagnosticar 6.2.5.2 Verificar e validar (V&V)	S – KPI - RG S – KPI - RG
6.2.6 Melhoria de processo e suporte	6.2.6.1 Implementando melhorias 6.2.6.2 Suporte de operações	S – KPI - RG S - KPI

FONTE: (adaptado de VAN ECK et al., 2015 tradução nossa)

## 6.2.1 Planejamento

### 6.2.1.1 Seleção de processos

Para a seleção da unidade para o estudo de caso dentro do CHC, inicialmente foi necessário identificar quais unidades estariam preparadas para a aplicação da ferramenta (mineração de processos), com relação ao registro de *logs* de eventos e características do processo. Para isso, durante 3 meses foram realizadas reuniões<sup>24</sup> com a diretoria do hospital para que os setores mais prováveis de estarem adequados fossem indicados. Foram elencados três setores: Farmácia, Unidade de Processamento de Materiais Esterilizados (UPME) e Unidade do Laboratório de Análises Clínicas (ULAC). Com os locais e o contato de cada um, foram realizadas reuniões com os responsáveis de cada um dos três setores para avaliar a viabilidade

<sup>24</sup> Formalmente foram realizadas três reuniões nessa etapa, duas reuniões na etapa de Mineração e Análise e uma nas demais, porém, ocorreram diversas trocas de informação por outros meios mais informais e difíceis de contabilizar, como por exemplo: e-mails, telefones, conversas nos corredores do hospital, dentre outros.

no que diz respeito à qualidade dos dados e ao processo em si. Após a apresentação da situação de cada setor, o setor de qualidade do hospital foi envolvido para estabelecer o alinhamento com os projetos já em andamento e selecionar um setor que fosse complementar ao programa de melhoria contínua. Dessa forma, foi selecionada a Unidade do Laboratório de Análises Clínicas.

Todos os contatos dos diferentes setores analisados, até que fosse feita a opção pela ULAC, possuem o cargo de chefe do respectivo setor. Além disso, sempre que adequado, foram apresentadas imagens com exemplos de resultados e tabelas contendo exemplos dos dados que deveriam ser extraídos do sistema.

Para a realização de projetos relacionados à saúde, foi necessário passar pelo Comitê de Ética em Pesquisa (CEP) da instituição. Foi então elaborado o projeto de pesquisa para aprovação pela ULAC e, em seguida, submetido à Plataforma Brasil<sup>25</sup> para apreciação do CEP-CHC. O período de tempo envolvido na elaboração do projeto, aprovação pela ULAC e aprovação pelo Comitê de Ética foi de sete meses. O projeto foi aprovado pelo Certificado de Apresentação para Apreciação Ética (CAAE) de nº 79033417.7.0000.0096.

#### 6.2.1.2 Identificação das questões de pesquisa

O projeto foi estruturado de modo a ter como questão de pesquisa inicial:

*“Como é o fluxo de exames na ULAC?”*

A opção por uma questão de pesquisa inicial mais abstrata vai ao encontro das sugestões da metodologia mencionadas na seção 5.1.1 *Process Mining Process Methodology* (PM<sup>2</sup>). Tal prática abre mais espaço para refinamento posterior através da análise exploratória, resultando em ideias concretas de melhoria e *insights* valiosos (VAN ECK et al., 2015).

---

<sup>25</sup> A Plataforma Brasil é uma base nacional e unificada de registros de pesquisas envolvendo seres humanos para todo o sistema CEP\*/Conep\*\*. Ela permite que as pesquisas sejam acompanhadas em seus diferentes estágios - desde sua submissão até a aprovação final pelo CEP e pela Conep, quando necessário - possibilitando inclusive o acompanhamento da fase de campo, o envio de relatórios parciais e dos relatórios finais das pesquisas (quando concluídas). <http://plataformabrasil.saude.gov.br/>

\* CEP - Comitês de Ética em Pesquisa

\*\* CONEP - Comissão Nacional de Ética em Pesquisa

### 6.2.1.3 Composição da equipe de projeto

Os profissionais envolvidos no projeto foram selecionados de modo a cobrir a esfera gerencial e técnica. Sendo assim, foram selecionados os seguintes cargos:

- i. **Gestora do negócio:** Chefe do laboratório;
- ii. **Especialista no negócio:** pessoa responsável pela gestão dos dados do laboratório (visão mais voltada ao processo, com formação na área da saúde);
- iii. **Especialista em sistemas:** pessoa responsável pela gestão do SIH (visão mais voltada aos dados, com formação na área de computação)
- iv. **Analista de processos:** o próprio pesquisador

Um ponto de atenção identificado na composição do time foi que apenas o Analista de Processos possuía conhecimento sobre mineração de processos. Com a seleção do processo, das questões de pesquisa e da equipe do projeto, é possível passar para a etapa de extração dos dados.

### 6.2.2 Extração

Nessa etapa, os dados disponibilizados pelo hospital foram advindos do Sistema de Informações Hospitalares (SIH) com a extração realizada pela responsável pela gestão de dados do laboratório (serviço de análises clínicas).

#### 6.2.2.1 Determinação do escopo

Para a determinação do escopo, foi discutido junto à gestora e à especialista do negócio quais os elementos estavam disponíveis no sistema e como seria interessante selecioná-los. Nesse ponto foi identificado um problema: o sistema não permite que haja grande manipulação de dados na extração propriamente dita, dessa forma, apenas o período de extração pôde ser parametrizado e todas as outras colunas foram extraídas para tratamento posterior. Foi determinado o período de 01/01/2016 a 31/12/2017 para ser extraído do sistema. Tal período foi selecionado pelo motivo descrito no próximo tópico.

#### 6.2.2.2 Extração de dados de eventos

Os dados são extraídos do sistema via geração de arquivos .csv; um por vez. Porém, foram identificados alguns pontos de atenção no processo de extração:

1) A versão disponível do editor de planilhas só tem a capacidade de armazenar 64k linhas por arquivo. Considerando que a ULAC processa exames de cerca de 500 pacientes diariamente e que pode haver diversos exames para cada requisição do paciente, o processo de extração fica prejudicado.

2) Não é possível visualizar quantas linhas serão importadas no período selecionado.

3) Quando são extraídas informações de um determinado período que ultrapassam as 64k linhas, o sistema não gera um aviso de que a quantidade de linhas excedeu o disponível no arquivo.

4) Para não sobrecarregar o sistema com o acesso aos dados, o setor em questão só pode realizar esse tipo de consulta e extração durante o período das 12h-13h de segunda a sexta. Esse período acaba sendo utilizado para consultas internas e toma grande parte dessa janela.

5) O tempo de geração do arquivo .csv, que varia conforme o volume de dados no arquivo, é um fator importante a ser considerado também, uma vez que a janela de tempo disponível (1h) é pequena.

6) Pelo fato de a solicitação ter sido feita no mês de dezembro, período em que as pessoas costumam tirar férias, a extração levou mais tempo.

Para contornar tais limitações, foram extraídos arquivos com diversos horizontes de tempo para delimitar um período seguro de extração (um período em que não houvesse linhas excedendo as 64k, que iriam desaparecer). Além disso, foi necessário que a responsável conseguisse diluir a extração ao longo de alguns dias para que fosse viável conciliar essa tarefa fora do planejado com os processos habituais da ULAC. Assim, foram extraídos 72 arquivos contendo 3.285.654 linhas no total. As 34 colunas resultantes da extração do SIH (Sistema de Informações Hospitalares) estão apresentadas na FIGURA 40.

FIGURA 40 - COLUNAS EXTRAÍDAS DO SIH

Nº	Coluna	Nº	Coluna
1	Requisicao	18	Hora_de_Entrada_Secao
2	Paciente	19	Quem_Recebeu_Amostra
3	Secao	20	Data_de_Liberacao
4	Codigo_da_Lotacao	21	Hora_de_Liberacao
5	Descricao_da_Lotacao	22	Quem_Liberou
6	Exame	23	Liberacao_Entrada_Seção_Dias
7	Nome_do_Exame	24	Liberacao_Entrada_Seção
8	Tipo_do_Requisitor	25	Liberacao_Coleta_Dias
9	Codigo	26	Liberacao_Coleta
10	Descricao_do_Requisitor	27	Liberacao_Requisicao_Dias
11	Data_da_Requisicao	28	Liberacao_Requisicao
12	Hora_da_Requisicao	29	Entrada_Seção_Coleta_Dias
13	Quem_Requisitou	30	Entrada_Seção_Coleta
14	Data_de_Coleta	31	Coleta_Requisição_Dias
15	Hora_da_Coleta	32	Coleta_Requisição
16	Quem_Coletou	33	Prazo_Urgente
17	Data_de_Entrada_Secao	34	Tipo_do_Prazo_1Hora_2Dia_3Semana_4Mes_5Ano_6Minuto

FONTE: O Autor (2018)

### 6.2.2.3 Transferência de conhecimento do processo

A transferência do conhecimento iniciou-se antes do processo de *Extração* propriamente dito e estendeu-se até a etapa de *Mineração e Análise*. Ela se deu inicialmente por meio de visitas *in loco* guiadas pela gestora do negócio com inclusão dos responsáveis por cada seção. Em seguida, a especialista no negócio e a especialista em sistemas tiveram participação direta em diversos momentos (por meio de reuniões e *e-mails*) para auxiliar na interpretação dos dados e sanar dúvidas pontuais de interpretação. Um ponto de atenção apresentado nessa etapa foi a necessidade de alinhamento dos termos utilizados por cada um dos envolvidos. Tanto do lado dos responsáveis pelo processo laboratorial quanto do analista de processos.

### 6.2.3 Processamento de dados

Essa é a etapa inicial do processo iterativo (Processamento de dados – Mineração e análise – Avaliação). Foram realizadas três iterações até obter o resultado final. De modo a facilitar a apresentação dos resultados, as iterações estão descritas e apenas o resultado oriundo da iteração final é apresentado nos subcapítulos referentes a cada etapa contida na iteração.



Nessa etapa de processamento de dados, o *feedback* da especialista do negócio com relação às conclusões intermediárias sobre os dados foi frequente (na maior parte das vezes, utilizando-se de imagens e desenhos). Com isso foi possível inicializar e atualizar o processamento de dados após cada iteração.

#### 6.2.3.1 Criação de visualizações

Antes de poder iniciar o processamento em si, os arquivos em .csv foram carregados via *Bulk Insert* em um banco de dados Microsoft SQL Server 2014 Express para poder relacionar os dados contidos em todos os 72 arquivos e facilitar o tratamento. Foram então identificadas as colunas que possuíam função meramente auxiliar (não iriam agregar informações na mineração) e removidas do conjunto de dados. Além disso, como as colunas relacionadas ao *timestamp* (Data\_da\_Requisicao, Hora\_da\_Requisicao, Data\_da\_Coleta, dentre outras) estavam separadas em data e hora, foram consolidadas em apenas uma. Por fim, para que fosse inserida a noção de *case*, foi necessário separar cada evento em uma linha distinta: ao invés de manter todas as atividades em uma mesma linha (FIGURA 41), foram colocadas em linhas distintas (FIGURA 42).

FIGURA 41 - SEM NOÇÃO DE CASE

Paciente X	Atividade 1	Timestamp 1	Atividade 2	Timestamp 2	Atividade 3	Timestamp 3	Atividade 4	Timestamp 4
------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------

FONTE: O Autor (2018)

FIGURA 42 - COM NOÇÃO DE CASE

Paciente X	Atividade 1	Timestamp 1
Paciente X	Atividade 2	Timestamp 2
Paciente X	Atividade 3	Timestamp 3
Paciente X	Atividade 4	Timestamp 4

FONTE: O Autor (2018)

Com essas alterações, o resultado das colunas encontra-se no QUADRO 7. As colunas presentes no QUADRO 7 foram utilizadas para inicializar o processo, porém, houve mudanças nas definições ao longo das iterações. O chaveamento (identificador)<sup>26</sup> iniciou-se apenas nas colunas “Requisicao” e “Paciente” conforme descrito, utilizando diversas colunas como atributo. Porém, posteriormente foi

<sup>26</sup> O chaveamento são os campos definidos como necessários para que uma instância configure um objeto único. No caso da presente dissertação, são as colunas necessárias para que não haja duplicatas no que diz respeito à requisições e pacientes.

identificada a necessidade de classificá-los como identificador (ID), conforme apresentado no QUADRO 8.

QUADRO 7 - COLUNAS UTILIZADAS PARA A PADRONIZAÇÃO DOS DADOS

<b>COLUNA</b>	<b>DESCRIÇÃO</b>
<b>Requisicao</b>	<i>Número da requisição feita pelo médico que pediu o exame</i>
<b>Paciente</b>	<i>Número de registro do paciente</i>
<b>Secao</b>	<i>Divisão da lotação em que o exame será processado (ex: sorologia, micologia, etc.)</i>
<b>Codigo_da_Lotacao</b>	<i>ID da lotação</i>
<b>Descricao_da_Lotacao</b>	<i>Divisão do hospital em que o exame será processado (ex: imunoquímica, bacteriologia, etc.)</i>
<b>Exame</b>	<i>ID do exame</i>
<b>Nome_do_Exame</b>	<i>Nome do exame que foi realizado</i>
<b>Tipo_do_Requisitor</b>	<i>Define a origem da requisição no nível macro (ex: ambulatório, convênio ou unidade de internação)</i>
<b>Codigo</b>	<i>ID do requisitor</i>
<b>Descricao_do_Requisitor</b>	<i>Define a origem da requisição no nível micro (ex: Unidade de internação de Neurocirurgia, Pediatria Emergência, etc.)</i>
<b>ATIVIDADE</b>	<i>Atividade que foi desempenhada</i>
<b>TIME_STAMP</b>	<i>Horário em que a atividade terminou</i>
<b>RESPONSAVEL</b>	<i>Pessoa responsável por apontar a atividade</i>

Fonte: O Autor (2018)

QUADRO 8 - CLASSIFICAÇÃO DAS COLUNAS EM CADA ITERAÇÃO

<b>COLUNA</b>	<b>1ª iteração</b>	<b>2ª iteração</b>	<b>3ª iteração</b>
<b>Requisicao</b>	ID	ID	ID
<b>Paciente</b>	ID	ID	ID
<b>Secao</b>	Atributo	ID	ID
<b>Codigo_da_Lotacao</b>	-	-	ID
<b>Descricao_da_Lotacao</b>	Atributo	ID	ID
<b>Exame</b>	-	-	-
<b>Nome_do_Exame</b>	Atributo	ID	ID
<b>Tipo_do_Requisitor</b>	Atributo	ID	ID
<b>Codigo</b>	-	-	-
<b>Descricao_do_Requisitor</b>	Atributo	ID	ID
<b>ATIVIDADE</b>	Atividade	Atividade	Atividade
<b>TIME_STAMP</b>	<i>Timestamp</i>	<i>Timestamp</i>	<i>Timestamp</i>
<b>RESPONSAVEL</b>	Recurso	Recurso	Recurso

FONTE: O Autor (2018)

Algumas colunas possuem valores que representam a mesma coisa, representando apenas o ID ou a descrição (ex: “Exame” e “Nome\_do\_Exame”). Um exemplo das colunas preenchidas encontra-se na FIGURA 43. Os nomes dos responsáveis pelo processamento das atividades foram anonimizados.

FIGURA 43 - EXEMPLO DE PREENCHIMENTO DAS COLUNAS

Requisicao	Paciente	Secao	Codigo_da_Lotacao	Descricao_da_Lotacao	Exame	Nome_do_Exame	Tipo_do_Requisitor	Codigo	Descricao_do_Requisitor	ATIVIDADE	TIME_STAMP	RESPONSAVEL
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	46	CITOLOGIA	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Requisicao	06/01/2016 10:29	Recurso 1
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	255	ALBUMINA	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Requisicao	06/01/2016 10:29	Recurso 1
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	288	LDH	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Requisicao	06/01/2016 10:29	Recurso 1
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	46	CITOLOGIA	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Coleta	06/01/2016 10:42	Recurso 2
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	255	ALBUMINA	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Coleta	06/01/2016 10:42	Recurso 2
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	288	LDH	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Coleta	06/01/2016 10:42	Recurso 2
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	46	CITOLOGIA	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Entrada_Secao	06/01/2016 10:42	Recurso 3
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	255	ALBUMINA	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Entrada_Secao	06/01/2016 17:27	Recurso 4
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	288	LDH	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Entrada_Secao	06/01/2016 17:27	Recurso 4
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	46	CITOLOGIA	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Liberacao	06/01/2016 17:22	Recurso 5
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	255	ALBUMINA	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Liberacao	06/01/2016 18:10	Recurso 6
160004681	23611570	2	573	SECAO DE IMUNOQUIMICA	288	LDH	UNIDADE DE INTERNAÇÃO	23	CENTRO DE TERAPIA SEMI INTENSIVA ADULTO	Liberacao	06/01/2016 18:10	Recurso 6

FONTE: O Autor (2018)

### 6.2.3.2 Agregação de eventos

Devido ao perfil dos dados disponibilizados, não foi necessário utilizar nem a agregação “is-a”, nem a “part-of”. No primeiro caso, devido à possibilidade de analisar cada tipo de exame processado individualmente (sem precisar agrupá-los). No segundo caso, a inexistência de subprocessos também fez com que esse tipo de agregação não fosse necessário.

### 6.2.3.3 Enriquecimento dos logs

Foi realizado o enriquecimento dos logs com a adição de dados externos (correlação). Essas informações foram obtidas após o fornecimento dos dados, com o uso de dados do RH por meio do contato com o setor de informática. A informação adicionada foi a unidade em que o funcionário está alocado (coluna [ALOCADO\_EM]). Um ponto de atenção nessa etapa foi que os dados fornecidos estavam relacionados à situação presente do quadro de funcionários. Portanto, funcionários que não se encontravam mais nas funções que desempenhavam quando foi feito o registro da atividade, podem não apresentar cadastro na unidade correta. O enriquecimento ocorreu apenas na terceira iteração, dessa forma, a sequência de alterações foi:

- **1ª Iteração:** sem enriquecimento
- **2ª Iteração:** sem enriquecimento
- **3ª Iteração:** Adição da coluna [ALOCADO\_EM] como atributo

### 6.2.3.4 Filtragem de logs

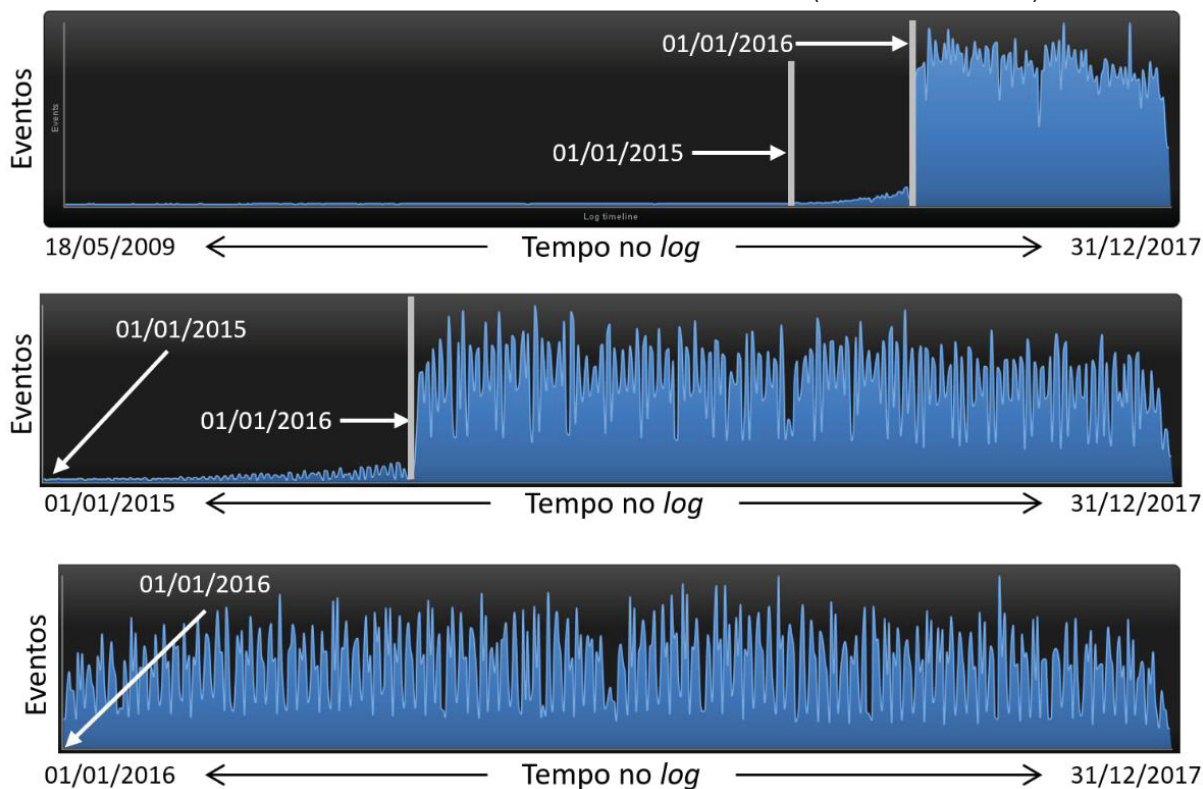
Para realizar a filtragem dos logs, os dados do banco de SQL foram salvos em formato .csv. Para a importação no Disco, a versão acadêmica possui uma limitação de 5M de eventos. Uma vez que o conjunto de dados obtido é maior do que o máximo

permitido nessa versão, foi feito contato com a Fluxicon (empresa criadora do Disco) para avaliar as possibilidades de obter uma licença comercial sem restrições. Como o objetivo do uso do *software* era exclusivamente a pesquisa, foi concedida uma licença com todas as funcionalidades disponíveis habilitadas, permitindo que o processamento do conjunto de dados fosse processado por inteiro.

Assim, os dados foram importados no Disco e as seguintes técnicas foram utilizadas:

- **Filtragem baseada em conformidade:** foram removidos *traces* incompletos, reduzindo em 1% o volume de dados para análise.
- **Slice and dice:** foi realizada a filtragem com relação à *timestamps* muito antigos (anteriores a 01/01/2016). A escolha do período considerado “antigo” se deu pela frequência de ocorrência dos *cases*, conforme ilustrado pela FIGURA 44. O primeiro gráfico mostra os dados após a “filtragem baseada em conformidade”. Ao longo das filtragens, o percentual analisado considerando os dados totais foi de 99% e 93% respectivamente.

FIGURA 44 - ETAPAS DE FILTRAGEM DOS DADOS (*SLICE AND DICE*)



FONTE: O Autor (2018)

Além disso, houve uma alteração nas estatísticas de cada log, conforme apresentado na TABELA 1.

TABELA 1 - ESTATÍSTICAS DOS LOGS AO LONGO DA FILTRAGEM

	Log inicial	Sem incompletos (FBC*)	01/2015-01/2017 (SAD**)	01/2016-01/2017 (SAD**)
<b>Eventos</b>	13.142.616	13.138.888	13.052.772	12.261.012
<b>Cases</b>	3.257.679	3.256.920	3.235.757	3.040.650
<b>Mediana da duração do case</b>	~36,4 horas	~36 horas	~34,3 horas	~31,9 horas
<b>Média da duração do case</b>	52,9 dias	52,9 dias	46,6 dias	35 dias

\*Filtragem baseada em conformidade

\*\**Slice and Dice*

FONTE: O Autor (2018)

O computador utilizado para processar o conjunto de dados foi um Intel Core i7-7500 @ 2.7GHz com 8,00 GB de memória RAM instalado. Foram alocados 6GB de memória para o processamento (o máximo recomendável para deixar o suficiente para uso do sistema). Mesmo assim, o conjunto de dados foi muito grande para o Disco definir a mediana exata, por isso os valores aproximados na terceira linha da TABELA 1.

## 6.2.4 Mineração e Análise

Enquanto as etapas anteriores tinham como preocupação maior a idealização do projeto, compreensão do processo e a preparação dos dados, nessa etapa foi onde a mineração de processos propriamente dita ocorreu. Foi dado ênfase ao uso de KPIs e representações gráficas ao longo das etapas, conforme discutido ao longo dos capítulos seguintes.

### 6.2.4.1 Descoberta de Processos

A descoberta do processo foi realizada com o auxílio do algoritmo de descoberta “Disco miner”, implementado em uma ferramenta comercial chamada Disco. Esse algoritmo é baseado no algoritmo “Fuzzy Miner” (GÜNTHER; VAN DER AALST, 2007) e ambos os algoritmos consideram a frequência dos eventos, bem

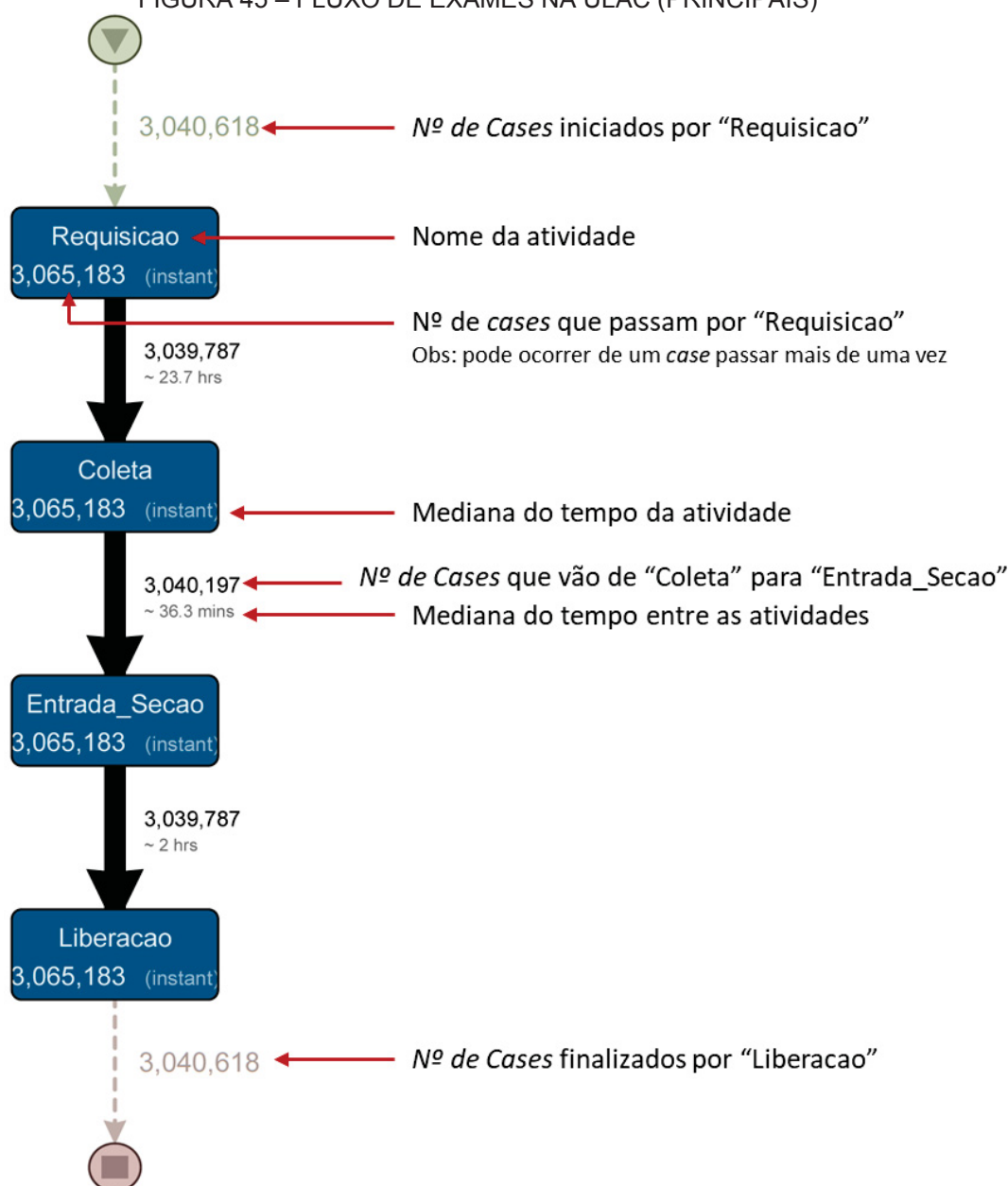
como sua sequência, buscando criar um modelo do processo (também conhecido como mapa do processo) (ALVAREZ et al., 2018).

As questões de pesquisa foram sendo aprimoradas com o passar das iterações para obter uma compreensão maior do processo, sua sucessão foi a seguinte:

– **1ª Iteração:** “Como é o fluxo de exames na ULAC?”

Para responder à primeira questão, foi elaborado um mapa do processo com auxílio do Disco com a explicação dos principais elementos (FIGURA 45). Para essa figura foi mantido apenas o fluxo de maior ocorrência no *log*.

FIGURA 45 – FLUXO DE EXAMES NA ULAC (PRINCIPAIS)

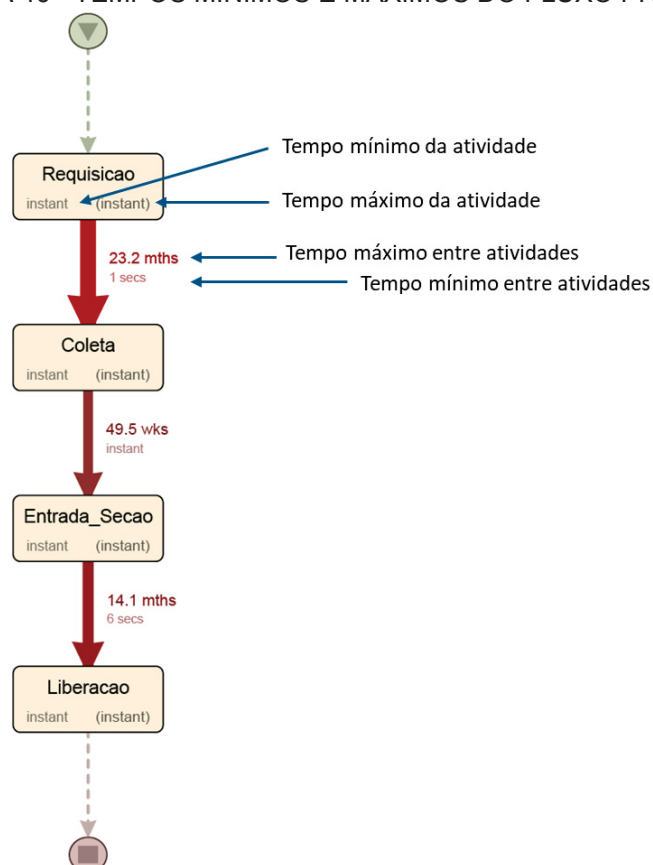


FONTE: O Autor (2018)

A partir da FIGURA 45 é possível observar que existem quatro atividades desenvolvidas dentro da ULAC: Requisição, Coleta, Entrada\_Seção e Liberação. O percentual de *cases* que passa por essas quatro atividades, nessa sequência, é de 99,9%. Os valores dentro de cada retângulo indicam o número de *cases* que passam pela atividade, assim, por exemplo: No caso da atividade “Requisição” existe o processamento de 3.065.253 *cases*. Isso constitui os 3.040.650 *cases* que entram por essa atividade e contabiliza os 24.603 que fazem o *loop* na atividade ( $3.065.253 - 24.603 = 3.040.650$ ). Da mesma forma, a saída são os 3.039.819 que vão para “Coleta” e os 831 que vão para a “Entrada\_Secao” ( $3.039.819 + 831 = 3.040.650$ ).

O rótulo “*instant*” ao lado do nome da atividade está apresentado pois só foi disponibilizado o *timestamp* do final da atividade. Caso fosse possível obter o *timestamp* do início, poderia ser feita a análise do tempo que a atividade em si leva. Os valores indicados abaixo do contador de *cases* são a mediana do tempo entre as atividades. A opção pela mediana ao invés da média foi feita devido aos valores muito altos de alguns *cases* particulares que distorcem a média. Os valores mínimos e máximos dos tempos entre atividades estão apresentados na FIGURA 46.

FIGURA 46 - TEMPOS MÍNIMOS E MÁXIMOS DO FLUXO PRINCIPAL



FONTE: O Autor (2018)



Analisando as figuras FIGURA 45 e FIGURA 46 em conjunto, é possível identificar que a discrepância do tempo mediano e do tempo máximo entre atividades é muito grande (em alguns casos, varia de 23 horas a 23 meses!). Uma possível explicação para tais discrepâncias é que na consulta realizada com o médico é realizada a Requisição tanto do exame relativo à atual consulta quanto do retorno (que justifica o a diferença de tempo).

Com o mapa do geral processo gerado, a próxima questão que surgiu diz respeito aos desvios com relação ao fluxo principal do processo.

- **2ª Iteração:** *“Existem desvios com relação ao processo descoberto? Quais os possíveis motivos?”*

Essa segunda parte, que tem a ver com conformidades, será respondida na próxima seção: *Verificação de conformidade*.

Por fim, em um estágio mais voltado à melhoria do processo (discutida com mais detalhes nas seções seguintes), foi levantada uma questão relativa aos pontos que são possíveis gargalos no processo.

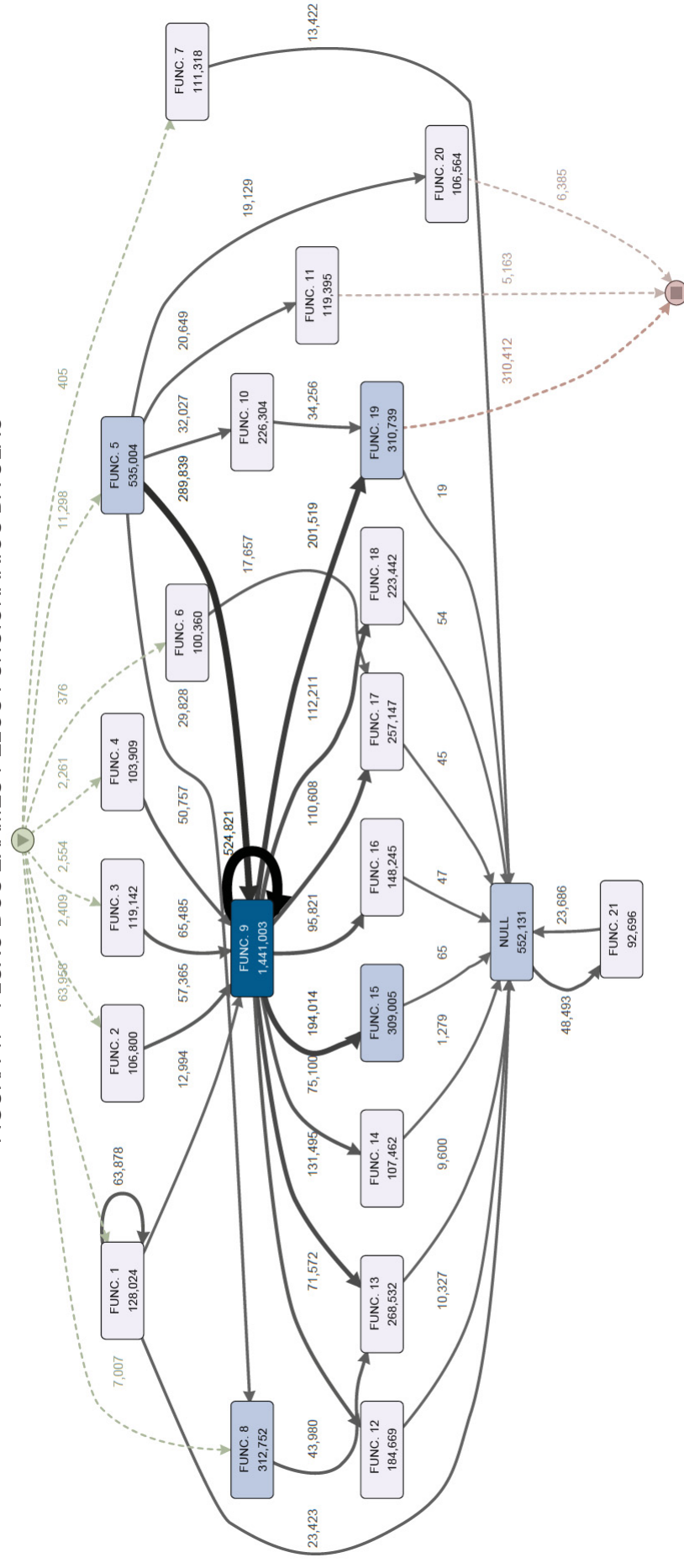
- **3ª Iteração:** *“Quais são os gargalos do processo?”*

Para responder a essa pergunta, foi alterada a coluna relativa à atividade presente no QUADRO 8 para que o RECURSO esteja mapeado para “Atividade”. Essa estratégia faz com que o mapa seja elaborado em função das pessoas que processam os eventos. A FIGURA 47 apresenta os recursos que processaram mais do que 307 atividades no período (sugestão do *software* para melhorar a apresentação gráfica). Os nomes relativos aos funcionários foram removidos e substituídos por “FUNC.” com números sequenciais.

A coloração da atividade indica a frequência de *cases* passando por ela. Assim, as atividades com azul mais escuro indicam maior frequência absoluta no processamento de *cases*. Dessa forma, o recurso “FUNC. 9” processou 1.441.003 *cases* no período considerado, seguido do recursos “FUNC. 5”, com 535.004. Em um primeiro momento, é possível concluir (erroneamente) que esses são gargalos do sistema. É nesse ponto em que o analista do negócio desempenha um papel crucial na interpretação dos dados.



FIGURA 47 - FLUXO DOS EXAMES PELOS FUNCIONÁRIOS DA ULAC



FONTE: O Autor (2018)

Os dois recursos mencionados apenas apontam (registram no sistema) o término do exame já processado, logo, são capazes de processar um volume muito mais alto do que os recursos que efetivamente realizam a coleta. Porém, com os dados fornecidos não é possível uma interpretação adequada dos gargalos do processo.

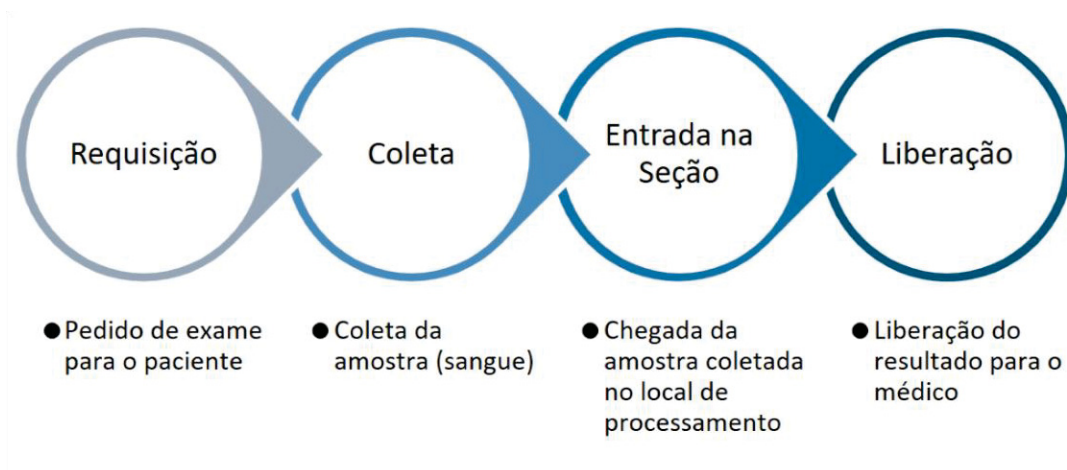
Os *timestamps* fornecidos contemplam apenas os tempos do fim da operação. Assim, não é possível mensurar o tempo efetivo empreendido em cada atividade por cada recurso, dificultando a identificação de gargalos. Em resumo, existem grandes espaços de tempo entre duas atividades identificadas e que seriam necessárias marcações de tempo para aprofundar a análise.

Além disso, existe um percentual das atividades (4,5%) que está passando pelo funcionário “NULL”, indicando que o exame foi cadastrado na atividade sem possuir um funcionário a ela. Esses eventos ocorrem na atividade “Entrada\_Secao” e foram causados por problemas de interfaceamento: o próprio equipamento dá entrada no exame e, por não ter um profissional cadastrado, não atrela o exame a um recurso

#### 6.2.4.2 Verificação de conformidade

Inicialmente, para desenvolver essa etapa, é necessário conhecer o processo concebido pela gestão da ULAC, apresentado na FIGURA 48.

FIGURA 48 - FLUXO DE EXAMES IDEALIZADO PELA ULAC



FONTE: O Autor (2018)

Os pacientes que realizam exames na ULAC já estão realizando algum tipo de tratamento no CHC. Assim, quando o médico precisa de algum exame para o diagnóstico, é feita uma “Requisição” de exame para o paciente. Uma requisição pode conter diversos exames para um mesmo paciente (por isso a necessidade de chaveamento nesses três campos e em outros mais). O *timestamp* correspondente refere-se ao momento de geração da requisição no sistema.

Com a requisição em mãos, o paciente vai até a ULAC para realizar a “Coleta” da amostra. Nos pontos de coleta do laboratório existe um dispositivo chamado de “robô” (um botão que é pressionado ao fim da extração da amostra), que registra o *timestamp* do fim da coleta. Os materiais coletados ficam em uma área de espera até serem levados ao local de processamento da amostra.

Uma vez que as amostras chegam ao laboratório em que serão processados (ex: imunoquímica, bacteriologia, etc.), a secretária registra o *timestamp* de entrada dos exames no sistema, registrando a “Entrada\_Secao”. Por fim, quando a amostra termina de ser processada e os resultados já estão disponíveis para o médico, é registrado o *timestamp* de “Liberacao”.

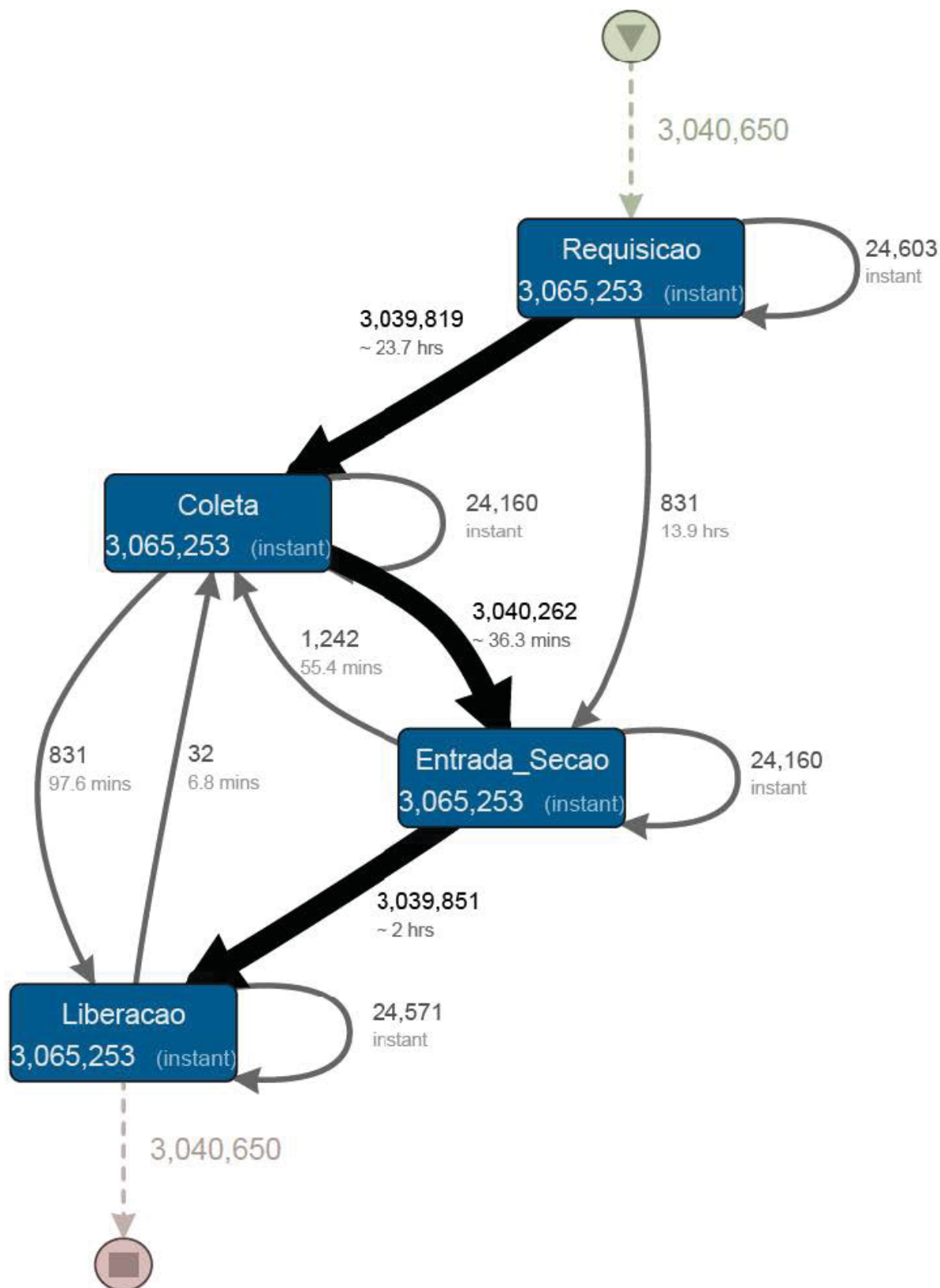
Expandindo a visualização dos caminhos seguidos pelos exames (ferramenta disponível no próprio Disco), é possível visualizar eventuais desvios ao que se tem planejado pela direção da ULAC. O mapa contendo esses desvios e o tempo mediano entre as atividades está apresentado na FIGURA 49.

Nesse novo mapa é possível identificar caminhos que fogem do que foi inicialmente concebido. Para compreender os motivos de tais desvios, foram consultados o gestor do negócio, o especialista do negócio e o especialista em sistemas. Após algumas reuniões, foi possível obter possíveis causas para tais desvios. O QUADRO 9 apresenta as combinações de atividades que não seguem a sequência esperada (Requisição, Coleta, Entrada\_Secao, Liberação), apresentadas na forma de atividade origem, atividade destino, número de ocorrências (segundo a FIGURA 49) e possíveis motivos para ocorrência.

Além dos casos apresentados no QUADRO 9, existem os casos em que o case faz um *loop* na própria atividade (ex: Requisição para Requisição). Uma possível explicação para esses casos é a falta de uma coluna considerada chave dentre as que foram fornecidas para a análise.

Na etapa de verificação de conformidade, o uso de KPIs é muito importante para medir o grau de conformidade entre processo desenhado e processo real. No caso da presente dissertação foram utilizados os Recursos, sendo a forma de contabilização quantos cases seguiram caminhos não esperados (2,936 cases, representando 0,097%). Uma vez que não foram encontrados na literatura valores de um percentual máximo de desvios para considerar um processo aderente ao modelo, o presente processo (com 99,903% de cases seguindo o fluxo desenhado), pode ser considerado aderente ao modelo desenhado.

FIGURA 49 - FLUXO COMPLETO DOS EXAMES NA ULAC



FONTE: O Autor (2018)

QUADRO 9 – FLUXOS ATÍPICOS REGISTRADOS NO LOG E POSSÍVEIS MOTIVOS

ORIGEM	DESTINO	OCORRÊNCIAS	POSSÍVEIS MOTIVOS
<i>Requisição</i>	<i>Entrada_Secao</i>	831	<ul style="list-style-type: none"> <li>– Algumas vezes, a coleta da amostra é realizada em um lugar diferente da área de coleta (ex: pacientes internados), fazendo com que não seja registrado o <i>timestamp</i> de “Coleta”</li> <li>– Quando a coleta é realizada em horários em que não há funcionários para operar o sistema.</li> <li>– Interfaceamento*</li> </ul>
<i>Coleta</i>	<i>Liberação</i>	831	<ul style="list-style-type: none"> <li>– Em alguns setores existem plantonistas que ainda não dominam completamente o sistema e, eventualmente, não registram a <i>Entrada_Secao</i>. Apenas a <i>Liberação</i> (situação atípica)</li> </ul>
<i>Entrada_Secao</i>	<i>Coleta</i>	1242	<ul style="list-style-type: none"> <li>– São requisições solicitadas pelo próprio pessoal do Laboratório para corrigir ou resolver algum problema de solicitação incorreta de exames (ex: alguma informação que faltou, que não está completa, alteração de exame no meio do caminho)</li> </ul>
<i>Liberação</i>	<i>Coleta</i>	32	<ul style="list-style-type: none"> <li>– O pedido é feito por telefone e não é gerada uma nova requisição (aproveita a antiga)</li> </ul>

\**Interfaceamento* - significa que quando ocorre a integração de uma máquina que processe um exame e automaticamente insira os dados no sistema, podem ocorrer erros e ajustes manuais até que a integração esteja completa.

FONTE: O Autor (2018)

Outros KPIs interessantes de incluir na análise seriam Custo e Qualidade. Porém, a massa de dados fornecida não forneceu informações suficientes para incluí-los na análise. O Tempo, apesar de ter sido fornecido (na forma de *timestamps*), não permite análises muito profundas, uma vez que não contempla o início e fim das atividade; os tempos analisados somam o tempo de espera com o tempo de processo, dificultando a interpretação correta do elemento tempo.

Em resumo, na etapa de *Verificação de conformidade* foram encontrados os seguintes pontos de atenção:

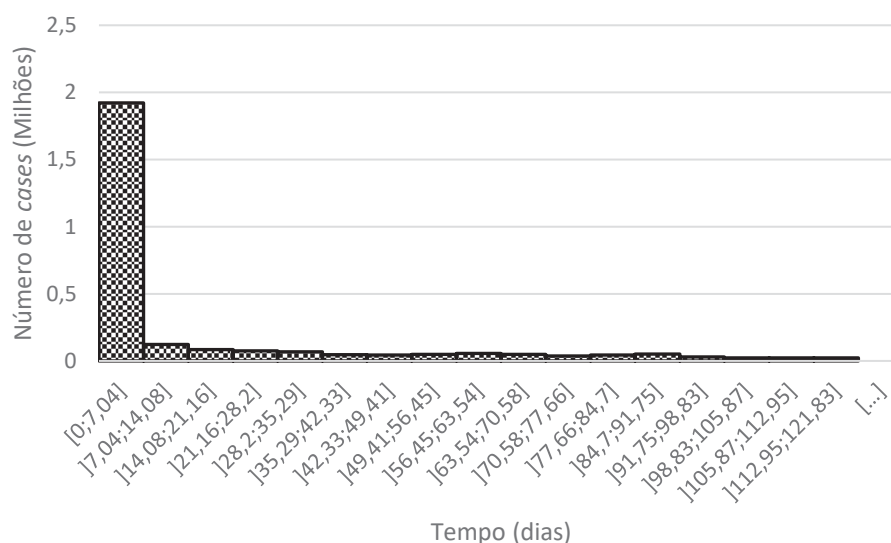
- Desvios com relação ao desenho do processo
- Falta de colunas para chaveamento (para remover o *loop* na própria atividade)
- Dificuldade de estabelecer os KPIs devido às colunas fornecidas
- Os *timestamps* fornecidos não apontam o início e fim das atividades, agrupando o tempo de atividades de espera e de processo efetivo.

#### 6.2.4.3 Análise do processo

Nessa etapa de Análise do processo foram exploradas informações adicionais que poderiam ser utilizadas para o aprimoramento do modelo.

A FIGURA 50 apresenta um histograma da duração que cada *case* leva para ser finalizado, sendo “dias” a unidade de medida de cada intervalo de classe. Só estão apresentadas as 17 primeiras classes para facilitar a visualização (o histograma completo possui 101 classes e as classes subsequentes possuem valores menores do que os das classes apresentadas).

FIGURA 50 – HISTOGRAMA DA DURAÇÃO DE CADA CASE (17 PRIMEIRAS CLASSES)



FONTE: O Autor (2018)

Com o histograma da FIGURA 50 é possível verificar que uma porção substancial dos *cases* é finalizado antes de 7,04 dias (63,2% dos 3,040,650 *cases*). O restante está diluído em dois grupos: (i) os 26,9% dos *cases* contidos nas 16 classes subsequentes (intervalo [7,04;121,83] ) e (ii) os 9,9% dos *cases* restantes, que encontram-se diluídos nas 84 classes restantes (intervalo [121,83;713] ).

Essa análise do tempo total do *case* contempla o tempo desde a geração da “Requisição” até a “Liberação” do exame para o médico. Dessa forma, é possível, por exemplo, que um *case* seja requisitado no dia 22/01/2016 e tenha as demais atividades realizadas no dia 28/12/2017 (intervalo real do *case* de maior duração). O intervalo entre as atividades de “Coleta”, “Entrada\_Secao” e “Liberação” não podem levar muito tempo entre si devido ao fato de ser material biológico e degradar com o tempo (variando de exame para exame).

#### 6.2.4.4 Aprimoramento

A etapa de aprimoramento pressupõe a apresentação de modelos de processo aprimorados, com base no que foi identificado nas etapas de *Descoberta do Processo*, *Verificação de Conformidade* e *Análise do Processo*. Porém, com uma aderência de 99,903% do processo com o modelo, não há indicativos de discrepância significativa entre os dois.

O que pode ser desenvolvido é um processo de adequação das diretrizes para fazer com que os tempos totais dos *cases* sejam reduzidos. O ponto mais relevante a explorar é o período entre a requisição e a coleta (variando de 1 segundo a 23,2 meses).

#### 6.2.5 Avaliação

Essa etapa é a última dentro da iteração e busca relacionar as conclusões da análise com as ideias de melhoria que vão ao encontro dos objetivos do projeto, buscando ideias de melhoria ou novas questões de pesquisa.

##### 6.2.5.1 Diagnosticar

O diagnóstico dos achados da mineração foram sendo desenvolvidos ao longo das etapas de *Processamento de Dados e Mineração e Análise* para manter a discussão próxima ao que foi identificado. Da mesma forma, a identificação/refino das questões de pesquisa foram feitas na seção 6.2.4.1 - *Descoberta de Processos*. Além disso, como não houve problemas identificados na etapa de descoberta, isso se estende à etapa de diagnóstico.



#### 6.2.5.2 Verificar e validar (V&V)

Da mesma maneira que o diagnóstico, a verificação e validação tiveram sua exploração próxima do ponto de menção. A discussão sobre causas-raiz subjacentes e projeção de ideias para possíveis melhorias no processo foi apresentada na seção 6.2.4.2 – *Verificação de Conformidade*. Assim, os pontos de atenção identificados anteriormente na etapa de verificação de conformidade, serão transportados para essa etapa de V&V.

#### 6.2.6 Melhoria de Processos e Suporte

Essa etapa final do PM2HC envolve a utilização das ideias de melhoria para modificar a execução real do processo.

##### 6.2.6.1 Implementar melhorias

A exploração e implementação das alternativas de melhoria, como o próprio modelo sugere, geralmente é um projeto separado. Para que eventuais modificações propostas sejam consideradas, há a necessidade de envolver outros setores dentro do Complexo Hospital de Clínicas. Isso envolve diversas atividades que fazem o fator tempo tornar-se crítico. Assim, estão elencados os pontos que justificam a não execução dessa etapa na ULAC dentro do presente projeto de mineração de processos:

- É necessário o estudo aprofundado das diretrizes de saúde para que as modificações estejam adequadas;
- Elaboração de um novo projeto e aprovação pelo Comitê de Ética;
- Busca de financiamento para o desenvolvimento das modificações;
- Burocracia inerente às alterações internas em instituições públicas;

##### 6.2.6.2 Suporte de operações

Da mesma forma que a implementação de melhorias, o suporte de operações, nesse contexto estudado, exigiria um projeto à parte. Os pontos necessários para seu desenvolvimento seriam:





Essa visão geral prioriza a visualização dos elementos que necessitaram de atenção/correção ao longo do projeto de mineração de processos da ULAC-CHC. Assim, pode-se dizer que a etapa de planejamento ocorreu sem maiores dificuldades, com a única dificuldade sendo o conhecimento de mineração de processos centralizado totalmente no analista de processo. As representações gráficas utilizadas junto aos *Stakeholders* permitiram o entendimento mais efetivo dos objetivos, da mesma forma que abreviou o processo de modo geral (comparando-se com a não utilização das representações gráficas).

O processo de extração dos dados apresentou uma série de dificuldades, em especial pela forma como ela ocorreu. Como os dados foram extraídos pela analista responsável pelo negócio (e não pela TI), o formato .csv tornou-se muito lento de extrair e tratar. Um limitante para que tivesse sido escolhida outra forma de extração foi o projeto submetido ao Comitê de Ética, que foi concebido de modo mais local (focando apenas a ULAC). Caso tivesse sido vislumbrada a possibilidade de extração via central de informática (junto ao analista de TI), possivelmente poderia ter sido obtido um arquivo .bak, que já contém todos os dados necessários e no formato correto. Precisando apenas tratá-los.

Outro ponto relacionado à extração diz respeito às informações necessárias para poder elaborar um *log* adequado à um projeto de mineração de processos. Com relação a isso, todos os campos puderam ser preenchidos satisfatoriamente. Porém, houve problemas com relação ao local de extração (foi recomendado o uso do sistema de faturamento). Essa informação só veio ao conhecimento do pesquisador após a finalização dos trâmites relativos ao Comitê de Ética (em torno de 11 meses entre o início/definição do que seria pesquisado até a aprovação)<sup>27</sup>, na etapa de revisão da literatura. Assim, os dados não puderam ser extraídos do sistema de faturamento.

A iteração ocorreu três vezes, porém nem todas as questões de pesquisa puderam ser respondidas, conforme explicado anteriormente.

A etapa de processamento dos dados ocorreu sem maiores dificuldades, sendo utilizados todos comentários extraídos da revisão da literatura: padrão do *timestamp*, criação de *sub-logs*, colunas necessárias em um *log* e criptografia dos dados. Além

---

<sup>27</sup> O tempo envolvido nas atividades com o Comitê de Ética estenderam-se dessa forma por não ser uma prática comum na área de engenharia a aprovação de projetos pelo comitê. Assim, foi necessário consultar uma vasta documentação para ser possível desenvolver um projeto cujas diretrizes éticas estejam de acordo com as boas práticas.

disso, a filtragem e a clusterização desempenharam um papel fundamental no processo de tratamento dos dados, facilitando a entrada de dados no Disco. Por fim, as representações gráficas, nessa etapa aproximaram-se de esquemas de tabelas (quais colunas estariam associadas ou poderiam ser enriquecidas) para definir o que seria relevante para cada etapa.

Na mineração e análise, a verificação de conformidade foi a etapa que apresentou maior número de dificuldades, em especial devido à qualidade dos dados. Isso dificultou o estabelecimento de métricas concretas para a análise do processo. Assim, apenas as representações gráficas (especialmente na etapa de verificação de conformidade para ilustrar os fluxos) puderam ser utilizadas. Os KPIs poderiam ser associados às representações gráficas, mas, infelizmente, isso não foi possível.

A etapa de avaliação foi sendo feita ao longo das iterações, assim, as únicas dificuldades encontradas relacionam-se à etapa de V&V (herdados da seção de verificação de conformidade). A carência de KPIs inviabilizou parcialmente a busca por ideias de melhoria. Assim, as melhorias sugeridas restringiram-se à maior atenção no cumprimento das regras pré-estabelecidas e atenção no processamento de exames em situações atípicas (ex: atividades realizadas por plantonistas, coletas fora do laboratório, equipamentos recém interfaceados, etc.).

Por fim, como já mencionado anteriormente, sugere-se um projeto separado para tratar da implantação das melhorias e do suporte das operações. Por esse motivo, não foi possível desenvolver essa etapa. Porém, foram levantadas as observações para que esse projeto à parte possa ser desenvolvido, que estão listadas na seção 6.2.6 Melhoria de Processos e Suporte.

#### 6.4 RESULTADO FINAL DO PM<sup>2</sup>HC

O estudo de caso na ULAC-CHC permitiu que a metodologia PM<sup>2</sup>HC pudesse ser aplicada e avaliada com dados extraídos de um processo real de saúde. A vantagem relativa do PM<sup>2</sup>HC sobre o PM<sup>2</sup>, no âmbito geral, é o maior detalhamento ao longo das etapas da metodologia. Tal nível de detalhamento faz que o projeto de mineração de processos seja implementado de forma mais efetiva e padronizada, permitindo que haja um processo evolutivo natural das métricas envolvidas e, conseqüentemente, de suas análises. Essa padronização do processo vai ao encontro

das demandas levantadas por diversos autores da área (FEI; MESKENS; MOREAU, 2009; FUNKNER; YAKOVLEV; KOVALCHUK, 2017; STAAL, 2010).

Em um nível mais específico, o PM<sup>2</sup>HC possui o envolvimento dos *Stakeholders* nas etapas iniciais do projeto e permite o delineamento dos objetivos de forma a atender a organização de modo mais efetivo (BETTENCOURT-SILVA et al., 2015; CHO et al., 2017; DELIAS et al., 2015; HUANG et al., 2015; MANS et al., 2009; PERIMAL-LEWIS et al., 2012; RAMOS, 2009; REBUGE; FERREIRA, 2012; RIEMERS, 2009). Enquanto que a presença dos *Stakeholders* nas etapas finais desempenha o papel de avaliar se tais objetivos foram atingidos, impactando diretamente na redução dos custos e na melhoria da qualidade do tratamento (YANG; SU, 2014). Esse envolvimento dos *Stakeholders* pode ser visto, inclusive, como uma forma de contribuir para a redução da distância existente entre o que é desenvolvido na academia e o que se passa fora dele (no caso, os ambientes de saúde). Com resultados mais palpáveis e replicáveis de forma mais consistente, pode vir a ser uma ferramenta adotada por instituições de saúde não apenas como ferramenta de pesquisa, mas como uma solução para algum problema (gargalos, tempo total de tratamento, utilização de ativos, dentre outros). Além disso, a tendência de crescimento da demanda por serviços na área de *healthcare* (MANS et al., 2009) torna ferramentas como o PM<sup>2</sup>HC cada vez mais importantes e necessárias no ambiente da saúde.

Outro diferencial do PM<sup>2</sup>HC é o uso de artifícios gráficos para auxiliar no entendimento das definições, ao longo praticamente todas as etapas do projeto. Tal prática permite que aspectos dúbios e/ou complexos possam ter um consenso entre as partes envolvidas de forma simples, conforme sugerem Meyer (1991) e Rivera & Duran (2004).

Com relação à clusterização e à filtragem, ambas já são mencionadas na metodologia PM<sup>2</sup>. Porém, não figuram como elementos determinantes na definição da complexidade do modelo extraído dos dados. No caso do PM<sup>2</sup>HC, é dada a ênfase nesses dois elementos, associando-os à etapa de *Processamento de Dados*.

A inserção dos KPIs, a partir da metade do projeto, torna-o muito mais robusto, uma vez que provê subsídios numéricos para a análise. Tal fato relaciona-se diretamente com a presença dos *Stakeholders*, permitindo o alinhamento dos resultados com a estratégia organizacional, conforme mencionado por Seify (2010).

Por fim, os comentários presentes no PM<sup>2</sup>HC dizem respeito à aspectos mais técnicos de cada etapa e trazem maior segurança para analistas de mineração de processos iniciantes, uma vez que fornecem diretrizes técnicas para o desenvolvimento de algumas etapas do projeto.

Ao longo do estudo de caso foram encontrados alguns pontos em que o PM<sup>2</sup>HC não pôde ser explorado adequadamente:

- Na etapa de definição das questões de pesquisa, foi dada abertura total para o desenvolvimento do projeto, dificultando a definição do objetivo do projeto. Dessa forma, o objetivo tornou-se um diagnóstico geral do laboratório.
- Os processos da ULAC-CHC estavam estruturados no formato de um modelo “lasanha” (com atividades e fluxo bem definido). Assim, não foi possível explorar como o PM<sup>2</sup>HC se comportaria no caso de um modelo “espaguete”.
- O enriquecimento dos *logs* não foi possível, devido ao perfil de controle dos dados na ULAC-CHC.
- O aprimoramento do modelo foi dificultado devido à falta de apontamentos do início das atividades (não permitindo identificar a sua duração).
- A etapa de Melhoria de Processos e Suporte não pôde ser explorada devido ao que foi apresentado no projeto submetido ao Comitê de Ética associado ao tempo disponível no mestrado e à burocracia envolvida em alterações em ambientes públicos.

Apesar dos pontos que não puderam ser explorados com o estudo de caso, o PM<sup>2</sup>HC comportou-se bem ao longo de seu desenvolvimento, guiando o percurso para obter o diagnóstico da ULAC-CHC, obter causas prováveis e ideias de melhoria.

Para que a aplicação do PM<sup>2</sup>HC possa ser potencializada, devem ser observados os elementos (que irão permear toda a aplicação), em especial na etapa de inicialização. Nessa etapa é definido o time do projeto e o seu objetivo. Caso esses dois elementos não estejam claros, a aplicação pode não alcançar o que foi estabelecido inicialmente e/ou demandar um maior número de iterações que, consequentemente, consome mais tempo dos envolvidos. Além disso, deve-se prestar especial atenção ao uso dos KPIs para munir os analistas e *stakeholders* de

ferramentas quantitativas para embasar uma tomada de decisão mais assertiva. Deve-se considerar um breve treinamento sobre a forma de processamento dos dados e etapas do projeto (atentando-se para as diferenças de formação dos membros do time).

Por fim, quando possível, adequar o sistema de informações para registrar as informações necessárias dentro do *log* e no formato adequado (*timestamp*).

## 6.5 APRENDIZADO COM A APLICAÇÃO DO PM<sup>2</sup>HC

Essa seção busca consolidar os achados da aplicação do PM<sup>2</sup>HC, compilando os achados ao longo da aplicação na ULAC, sob a perspectiva do local de aplicação.

### 6.5.1 Compilação dos achados

Inicialmente, houve a necessidade de ambientar-se com as práticas do comitê de ética, uma vez que qualquer alteração do projeto necessita aprovação do conselho. Esse ponto fez com que o cronograma inicial tivesse que ser alterado em alguns pontos ao longo do projeto. Dessa forma, recomenda-se que, quando for necessária a apreciação ética do projeto, seja feito contato com os responsáveis para que o projeto seja estruturado da maneira mais adequada possível, de modo a reduzir a necessidade de emendas.

Já com relação ao projeto, os elementos iniciais do PM<sup>2</sup>HC (*Stakeholders*, Representações Gráficas, Clusterização, Filtragem e *Key Performance Indicators*) permitem um levantamento do que é interessante em um projeto de mineração de processos, conferindo subsídios para eventuais melhorias no processo da ULAC. Da mesma forma, os comentários auxiliam os envolvidos mais ligados com a área técnica de processamento de dados.

O processo de transferência do conhecimento entre as funcionárias da ULAC e o analista de processos permitiu que fosse feita uma reflexão sobre a maneira como os processos vem sendo efetivamente desenvolvidos. Essa reflexão permitiu um maior conhecimento do processo desenvolvido dentro das atribuições de cada um dos funcionários assim como a troca de informações sobre situações de como algo que ocorre em um setor tem impacto em outros setores.

A partir dessa reflexão, o mapa do processo (apresentando o fluxo de exames na ULAC) pôde ser interpretado mais facilmente, ilustrando de forma gráfica informações relativas ao volume de processamento de exames e tempo de atividades. Além disso, foi possível contabilizar o percentual de aderência ao modelo do processo idealizado (99,903%). Da mesma forma, foi possível identificar que alguns casos apresentam discrepâncias entre o seu tempo mediano e tempo máximo entre atividades (em alguns casos, varia de 23 horas a 23 meses). Além da necessidade de incorporar o *timestamp* do início da atividade nos registros para uma análise mais rica.

No que diz respeito ao fluxo de exames pelos funcionários da ULAC, houve a identificação do volume de processamento individualizado. Essa informação permite que, em uma eventual modificação dos processos, haja a realocação de pessoal para pontos de maior necessidade. Ainda nesse tópico, foi identificado que 4,5% das atividades estão sem o funcionário responsável cadastrado devido a problemas de interfaceamento, colocando um ponto de atenção quando da integração de novos equipamentos ao sistema.

Na verificação de conformidade do fluxo (confrontar o que efetivamente ocorre com o processo projetado), foi apresentado graficamente o volume de desvios e em que pontos ocorrem. Além disso, por meio da transferência de conhecimento realizada entre os integrantes do time, foi possível elencar suas potenciais causas-raiz.

Foi apresentada uma forma alternativa de visualização da duração de cada *case*; um histograma. Nele foi possível identificar qual percentual dos *cases* tem seu tempo total contemplado dentro de cada classe do histograma.

Por fim, os desvios e anormalidades identificados permitem que haja a adequação pontual de alguns processos para que o processamento dos exames seja melhorado. Essa melhoria pode vir de diversas formas: padronização de processos, capacitação das pessoas envolvidas (em especial quando há trocas de função ou alta rotatividade), incremento nas formas de controle, dentre outras.



## 7 CONSIDERAÇÕES FINAIS

Na presente dissertação foi apresentado um estudo relativo à uma metodologia de mineração de processos aplicada à saúde. A motivação inicial para o trabalho foi a dificuldade de localizar na literatura metodologias de implementação de projetos desse tipo, associado ao seu retorno potencial para a instituição. Partindo dessa carência metodológica, foram definidos objetivo geral e objetivos específicos, que puderam ser atingidos por meio de levantamentos bibliográficos e uma aplicação na Unidade do Laboratório de Análises Clínicas do Complexo Hospital de Clínicas da Universidade Federal do Paraná.

O modelo base selecionado (VAN ECK et al., 2015) atendeu o propósito da pesquisa, uma vez que permitiu a incorporação dos complementos levantados ao longo das revisões da literatura (Representações Gráficas, *Stakeholders*, Clusterização, Filtragem, KPIs e os comentários) para tornar-se PM<sup>2</sup>HC.

Assim, quando foi realizada a aplicação do PM<sup>2</sup>HC na ULAC-CHC, foi possível utilizá-lo sem modificações, em especial, se comparado com o PM<sup>2</sup>, onde muitas das etapas não estavam descritas em um nível de detalhe que permitisse o desenvolvimento adequado do projeto. Nesse ponto, a aplicação do PM<sup>2</sup>HC possibilitou a redução/eliminação das dúvidas que surgiram ao longo desse processo.

Ao longo da aplicação do PM<sup>2</sup>HC foram encontradas dificuldades que tornaram-se sugestões de melhoria para o local de aplicação (como, por exemplo, os *timestamps* que são gravados apenas no fim da operação ou a necessidade de maior atenção no interfaceamento). Outra questão, relacionada à aplicação, foi o fato de não ter sido possível desenvolver/implementar as ideias de melhoria, requerendo um projeto à parte. Porém, com o diagnóstico apresentado na presente pesquisa, tal projeto pode ser abreviado e será direcionado aos pontos levantados no diagnóstico. Dessa forma, é possível considerar que o PM<sup>2</sup>HC foi efetivo nas etapas compreendidas do Planejamento à Avaliação (etapa final da iteração).

Além disso, considerando a forma de como a metodologia está estruturada, associada aos achados no estudo de caso, é possível dizer que o PM<sup>2</sup>HC consegue suprir as necessidades do projeto de modo a fechar o ciclo completo da metodologia, caso a etapa de Melhoria de Processo e Suporte seja incorporada a um único projeto.

Como trabalhos futuros, a metodologia *L\* lifecycle model* pode ser testada no lugar do PM<sup>2</sup>, uma vez que, após a análise dos dados, identificou-se que o modelo da



ULAC não era do tipo espaguete (critério de exclusão do *L\* lifecycle model*, que lida melhor com modelos do tipo “lasanha”). Para efeito de comparação com o PM<sup>2</sup>HC, sugere-se que esse teste seja realizado em setores que desenvolvam um número baixo de atividades (próximo ao que foi estudado na presente dissertação) e que não haja muitos *cases* que fujam do fluxo principal. Outra opção de trabalhos futuros, seria a aplicação do PM<sup>2</sup>HC em um ambiente com um modelo de processo menos padronizado, buscando confrontá-lo com os resultados obtidos no estudo de caso apresentado.

## REFERÊNCIAS

VAN DER AALST, W. M. P. Process discovery: Capturing the invisible. **IEEE Computational Intelligence Magazine**, v. 5, n. 1, p. 28–41, 2010.

VAN DER AALST, W. M. P. **Process mining: discovery, conformance and enhancement of business processes**. Springer, 2011.

VAN DER AALST, W. M. P. **Process Mining: Data Science in Action**. Second ed. Springer, 2013a.

VAN DER AALST, W. M. P. **Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining**. 2013b.

VAN DER AALST, W. M. P. **Process Mining: Data Science in Action**. Springer, 2016.

VAN DER AALST, W. M. P.; ADRIANSYAH, A.; DE MEDEIROS, A. K. A.; ARCIERI, F.; BAIER, T.; BLICKLE, T.; BOSE, J.C.; et al. Process mining manifesto. **Lecture Notes in Business Information Processing**, v. 99 LNBIP, n. PART 1, p. 169–194, 2012.

VAN DER AALST, W. M. P.; HOFSTEDE, A. H. M.; WESKE, M. Business Process Management: A Survey. 1st International Conference, BPM 2003 Eindhoven, The Netherlands, June 26-27, 2003, Proceedings. **Anais...** . p.1–12, 2003. Springer.

VAN DER AALST, W. M. P.; WEIJTERS, A. J. M. M. Process mining: a research agenda. **Computers in Industry**, v. 53, n. 3, p. 231–244, 2004. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0166361503001945>>. Acesso em 27 setembro 2017

ABPMP. **BPM CBOK - Versão 3.0**. ABPMP Brasil, 2013.

ALVAREZ, C.; ROJAS, E.; ARIAS, M.; Munoz-gama, J.; Sepúlveda, M.; Herskovic, V.; Capurro, D. Discovering role interaction models in the Emergency Room using Process Mining. **Journal of Biomedical Informatics**, v. 78, p. 60–77, 2018. Elsevier. <<https://doi.org/10.1016/j.jbi.2017.12.015>>.

BAKER, K.; DUNWOODIE, E.; JONES, R. G.; NEWSHAM, A.; JOHNSON, O.; PRICE, C. P.; WOLSTENHOLME, J. Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy. **International Journal of Medical Informatics**, v. 103, p. 32–41, 2017. Elsevier Ireland Ltd. <<http://dx.doi.org/10.1016/j.ijmedinf.2017.03.011>>.

BEIMBORN, D.; JOACHIM, N. The joint impact of service-oriented architectures and business process management on business process quality: An empirical evaluation and comparison. **Information Systems and e-Business Management**, v. 9, n. 3, p. 333–362, 2011.

BETTENCOURT-SILVA, J. H.; CLARK, J.; COOPER, C. S.; MILLS, R.; RAYWARD-SMITH, V. J.; de la IGLESIA, B. Building Data-Driven Pathways from Routinely Collected Hospital Data: A Case Study on Prostate Cancer. **JMIR Medical Informatics**, v. 3, n. 3, p. e26, 2015. Disponível em: <<http://medinform.jmir.org/2015/3/e26/>>. Acesso em: 31 janeiro 2018

BIUK-AGHAI, R. P.; SIMOFF, S. J. An Integrative Framework for Knowledge Extraction in Collaborative Virtual Environments. GROUP '01 Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work. **Anais...** p.61–70, 2001. Boulder, USA.

BIUK-AGHAI, R. P.; SIMOFF, S. J.; DEBENHAM, J. From Ad-hoc to Engineered Collaboration in Virtual Workspaces. Proceedings of the Eleventh Americas Conference on Information Systems. **Anais...** 2005. Omaha, EUA.

BLUM, T.; PADOY, N.; FEUSSNER, H.; NAVAB, N. Workflow mining for visualization and analysis of surgeries. **International Journal of Computer Assisted Radiology and Surgery**, v. 3, p. 379–386, 2008. Disponível em: <[http://download.springer.com/static/pdf/828/art%253A10.1007%252Fs11548-008-0239-0.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs11548-008-0239-0&token2=exp=1496172391~acl=%2Fstatic%2Fpdf%2F828%2Fart%25253A10.1007%25252Fs11548-008-0239-0.pdf](http://download.springer.com/static/pdf/828/art%253A10.1007%252Fs11548-008-0239-0.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs11548-008-0239-0&token2=exp=1496172391~acl=%2Fstatic%2Fpdf%2F828%2Fart%25253A10.1007%25252Fs11548-008-0239-0.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs11548-008-0239-0&token2=exp=1496172391~acl=%2Fstatic%2Fpdf%2F828%2Fart%25253A10.1007%25252Fs11548-008-0239-0.pdf)>. Acesso em: 30 maio 2017.

BOSE, R. P. J. C.; VAN DER AALST, W. M. P. Abstractions in process mining: A taxonomy of patterns. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 5701 LNCS, p. 159–175, 2009.

BOSE, R. P. J. C.; MANS, R. S.; VAN DER AALST, W. M. P. Wanna Improve Process Mining Results? It's High Time We Consider Data Quality Issues Seriously. **IEEE Symposium on Computational Intelligence and Data Mining (CIDM)**, p. 127–134, 2013. Disponível em: <<http://www.tue.nl/en/publication/ep/p/d/ep-uid/280745/>>. Acesso em: 09 jan. 2018

BOZKAYA, M.; GABRIELS, J.; WERF, J. M. VAN DER. Process Diagnostics: A Method Based on Process Mining. International Conference on Information, Process, and Knowledge Management. **Anais...** p.22–27, 2009. IEEE computer society. Disponível em: <<http://ieeexplore.ieee.org/document/4782560/>>. Acesso em: 11 mar. 2018

BRASIL. Ministério da Educação. Hospitais universitários. Brasília. 2012 Disponível em: <[http://portal.mec.gov.br/index.php?Itemid=512&id=12267&option=com\\_content&view=article](http://portal.mec.gov.br/index.php?Itemid=512&id=12267&option=com_content&view=article)>. Acesso em: 20 dez 2018.

BRASIL. MINISTÉRIO DA SAÚDE. Ministro vai usar tecnologia para melhorar gestão do SUS. Disponível em: <<http://www.brasil.gov.br/saude/2016/05/ministro-vai-usar-tecnologia-para-melhorar-gestao-do-sus>>. Acesso em: 30 set. 2017.

VOM BROCKE, J.; SCHMIEDEL, T.; RECKER, J.; TRKMAN, P.; MERTENS, W.; VIAENE, S. Ten principles of good business process management. **Business Process Management Journal**, v. 20, n. 4, p. 530–548, 2014. Disponível em: <<http://www.emeraldinsight.com/doi/10.1108/BPMJ-06-2013-0074>>. 28 out. 2018

CAVERLEE, J.; BAE, J.; WU, Q.; Liu, L.; Pu, C.; Rouse, W. B. Workflow management for enterprise transformation. **Information Knowledge Systems Management**, v. 6, n. 1/2, p. 61–80, 2007. Disponível em: <<http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=25215507&lang=de&site=ehost-live>>. Acesso em 28 out. 2018

CHALMERS, A. F. **O QUE É CIÊNCIA AFINAL?**. Editora brasiliense, 1993.

CHILDE, S. J.; WEAVER, A. M.; MAULL, R. S.; SMART, P. A.; BENNETT, J. The application of generic process models in business process re-engineering. Proceedings of the IFIP TC5/WG5.7 Working Conference on Re-engineering the Enterprise. **Anais...** p.110–120, 1995. Boston: Springer. Disponível em: <[https://link.springer.com/content/pdf/10.1007%2F978-0-387-34876-6\\_11.pdf](https://link.springer.com/content/pdf/10.1007%2F978-0-387-34876-6_11.pdf)>. Acesso em 28 out. 2018

CHO, M.; SONG, M.; COMUZZI, M.; YOO, S. Evaluating the effect of best practices for business process redesign: An evidence-based approach based on process mining techniques. **Decision Support Systems**, v. 104, p. 92–103, 2017. Elsevier B.V. <<https://doi.org/10.1016/j.dss.2017.10.004>>.

COOK, J. E.; WOLF, A. L. Discovering models of software processes from event-based data. **ACM Transactions on Software Engineering and Methodology**, v. 7, n. 3, p. 215–249, 1998.

CRUZ, J. I. B. **CONTRIBUIÇÕES NA AVALIAÇÃO DE CONFORMIDADE DE PROCESSOS DE DESENVOLVIMENTO DE SOFTWARE POR MINERAÇÃO DE PROCESSOS**, 2010. PUC-RS.

DAGLIATI, A.; SACCHI, L.; ZAMBELLI, A.; Tibollo, V.; Pavesi, L.; Holmes, J.H; Bellazzi, R. Temporal electronic phenotyping by mining careflows of breast cancer patients. **Journal of Biomedical Informatics**, v. 66, p. 136–147, 2017. Elsevier Inc. <<http://dx.doi.org/10.1016/j.jbi.2016.12.012>>.

DAMOUS, I.; ERLICH, H. O ambulatório de saúde mental na rede de atenção psicossocial: reflexões sobre a clínica e a expansão das políticas de atenção primária. **Physis Revista de Saúde Coletiva** v.27 n. 4 p. 911-932, 2017. <<http://dx.doi.org/10.1590/S0103-73312017000400004>>

DAVENPORT, T. H. **Process Innovation: Reengineering Work Through Information Technology**. Harvard Business School Press, 1993.

DEFOSSEZ, G.; ROLLET, A.; DAMERON, O.; INGRAND, P. Temporal representation of care trajectories of cancer patients using data from a regional information system: An application in breast cancer. **BMC Medical Informatics and Decision Making**, v. 14, n. 1, p. 1–15, 2014.

DELIAS, P.; DOUMPOS, M.; GRIGOROUDIS, E.; MANOLITZAS, P.; MATSATSINIS, N. Supporting healthcare management decisions via robust clustering of event logs. **Knowledge-Based Systems**, v. 84, p. 203–213, 2015. Elsevier B.V. <<http://dx.doi.org/10.1016/j.knosys.2015.04.012>>. .

DRESCH, A. **Design Science e Design Science Research como Artefatos Metodológicos para Engenharia de Produção**, 2013. Departamento de Engenharia de Produção e Sistemas. Unisinos. Disponível em: <<http://www.repositorio.jesuita.org.br/handle/UNISINOS/4075>>. Acesso em: 07 jan. 2018

DUNCAN, R. S.; CHRISTOPHER, P. H.; PETER, K.; BOB, S.; BRIAN, W. Elements of a business process management system: theory and practice. **Business Process Management Journal**, v. 13, p. 91–107, 2007. <<http://dx.doi.org.proxy1-bib.sdu.dk:2048/10.1108/14637150710721140>>. .

EARL, M. J. The new and the old of business process redesign. **The Journal of Strategic Information Systems**, v. 3, n. 1, p. 5–22, 1994. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0963868794900035>>. Acesso em: 28 out. 2017

EBSERH. EMPRESA BRASILEIRA DE SERVIÇOS HOSPITALARES. HC - Hospital de Clínicas. Disponível em: <<http://www.ebserh.gov.br/web/chc-ufpr/hc>>. Acesso em: 08 dez. 2018.

EBSERH. EMPRESA BRASILEIRA DE SERVIÇOS HOSPITALARES. Linha do tempo - HC/UFPR. Disponível em: <<http://www.ebserh.gov.br/web/chc-ufpr/linha-do-tempo>>. Acesso em: 08 dez. 2018.

VAN ECK, M. L.; LU, X.; LEEMANS, S. J. J.; VAN DER AALST, W. M. P. PM2: A process mining project methodology. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 9097, p. 297–313, 2015.

EPURE, E. V; INGVALDSEN, J. E.; DENECKERE, R.; SALINESI, C. Process Mining for Recommender Strategies Support in News Media. 2016 IEEE Tenth International Conference on Research Challenges in Information Science (RCIS). **Anais...** 2016.

ESFAHANI, F. S.; MURAD, M. A. A.; SULAIMAN, M. N. B.; UDZIR, N. I. Adaptable Decentralized Service Oriented Architecture. **The Journal of Systems & Software**, v. 84, n. 10, p. 1591–1617, 2011. Elsevier Inc. <<http://dx.doi.org/10.1016/j.jss.2011.03.031>>.

FANG, X.; JIANG, C.; YIN, Z.; FAN, X. The Trustworthiness Analyzing of Interacting Business Process Based on the Induction Information. **Computer Science and Information Systems**, v. 8, n. 3, 2011.

FEI, H.; MESKENS, N.; MOREAU, C.-H. Clustering of Patients' Trajectories with an Auto-Stopped Bisecting K-Medoids Algorithm. Proceedings of the 13th IFAC

Symposium on Information Control Problems in Manufacturing. **Anais...** . v. 12, p.135–154, 2009. Moscow, Russia: IFAC.

FERNANDEZ-LLATAS, C.; LIZONDO, A.; MONTON, E.; BENEDI, J.; TRAVER, V. Process Mining Methodology for Health Process Tracking Using Real-Time Indoor Location Systems. **Sensors**, , n. Cmmi, p. 29821–29840, 2015.

FUNKNER, A. A.; YAKOVLEV, A. N.; KOVALCHUK, S. V. Data-driven modeling of clinical pathways using electronic health records. *Procedia Computer Science*. **Anais...** . v. 121, p.835–842, 2017. Barcelona, Spain: Elsevier B.V. <<https://doi.org/10.1016/j.procs.2017.11.108>>. .

GANDULFO, P. I. **Método de Mineração de Processos para Auxílio à Tomada de Decisão: um Estudo de Caso no controle de férias**, 2016. Universidade de Brasília. Disponível em: <<http://repositorio.unb.br/handle/10482/22059>>. Acesso em: 23 jun. 2018

GERHARDT, R.; VALIATI, J. F.; SANTOS, J. V. C. DOS. An Investigation to Identify Factors that Lead to Delay in Healthcare Reimbursement Process: A Brazilian case. **Big Data Research**, v. Article in, p. 1–10, 2018. Elsevier Inc. <<https://doi.org/10.1016/j.bdr.2018.02.006>>.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 2008.

GLOBERSON, S.; ZWIKAEI, O. The Impact of the Project Manager on Project Management Planning Processes. **Project Management Journal**, v. 33, n. 3, p. 58–64, 2002. Disponível em: <<https://journals.ala.org/index.php/rusq/article/view/5793>>. Acesso em: 09 jan. 2018

GRAY, J. Data Management: Past, Present, and Future. **IEEE Computer** **29**, v. 29, n. June, p. 38–46, 2007. Disponível em: <<http://arxiv.org/abs/cs/0701156>>. 28 out. 2017

GÜNTHER, C. W.; VAN DER AALST, W. M. P. Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. *Lecture Notes in Computer Science*. **Anais...** . v. 4714, p.328–343, 2007. Berlin: Springer. Disponível em: <[http://link.springer.com/10.1007/978-3-540-75183-0\\_24](http://link.springer.com/10.1007/978-3-540-75183-0_24)>. 09 jan. 2018

GÜNTHER, C. W.; ROZINAT, A.; VAN DER AALST, W. M. P.; Van, K.; Uden Cardio. **Monitoring Deployed Application Usage with Process Mining**. Eindhoven, 2008.

GUPTA, S. **Workflow and Process Mining in Healthcare (MSc thesis)**, 2007. Technische Universiteit Eindhoven. Disponível em: <<http://alexandria.tue.nl/extra2/afstversl/wsk-i/gupta2007.pdf>>. Acesso em: 30 maio 2017.

HARMON, P. **A Guide for Business Managers and BPM and Six Sigma Professionals**. Morgan Kaufmann, 2007.



VAN DER HEIJDEN, T. H. C. **Process Mining Project Methodology: Developing a General Approach to Apply Process Mining in Practice (Msc. Thesis)**, 2011. Eindhoven University of Technology.

HERBST, J. A Machine Learning Approach to Workflow Management. **Machine Learning: ECML 2000**, v. 1810, p. 183–194, 2000. Disponível em: <<http://www.springerlink.com/index/m7t4x71227vm1w43.pdf>>. Acesso em: 18 mar. 2018

HERDER, P. M.; WEIJNEN, M. P. C. Knowledge management to improve the quality of process design and the design process. **Computers and Chemical Engineering**, v. 23, n. SUPPL. 1, p. S815–S818, 1999. Elsevier Science Ltd. <[http://dx.doi.org/10.1016/S0098-1354\(99\)80200-2](http://dx.doi.org/10.1016/S0098-1354(99)80200-2)>.

HUANG, Z.; DONG, W.; BATH, P.; JI, L.; DUAN, H. On mining latent treatment patterns from electronic medical records. **Data Mining and Knowledge Discovery**, v. 29, n. 4, p. 914–949, 2015. Springer US. <<http://dx.doi.org/10.1007/s10618-014-0381-y>>.

INSPER. **O panorama da saúde no Brasil**. 2012.

JUHANAK, L.; ZOUNEK, J.; ROHLÍKOVÁ, L. Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. **Computers in Human Behavior**, v. Article in, 2017.

KHOWAJA, A. R. Process Mining Techniques: An Application to Time Management. Ninth International Conference on Graphic and Image Processing (ICGIP 2017). **Anais...**, 2018.

KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3**. Durhan, UK, 2007.

KURNIATI, A. P.; JOHNSON, O.; HOGG, D.; HALL, G. Process mining in oncology: A literature review. **Proceedings of the 6th International Conference on Information Communication and Management, ICICM 2016**, , n. i, p. 291–297, 2016.

LANG, M.; BÜRKLE, T.; LAUMANN, S.; PROKOSCH, H.-U. Process Mining for Clinical Workflows: Challenges and Current Limitations. In: M. C. Andersen, S.K., Klein, M.C.A., Schulz, S., Aarts, J., Mazzoleni (Org.); **eHealth Beyond the Horizon – Get IT There**. 136<sup>o</sup> ed, p.229–34, 2008. Amsterdam: IOS Press. Disponível em: <<https://ai2-s2-pdfs.s3.amazonaws.com/cb8d/d9182c568638cff00b7370c6d6edc1b392f3.pdf>>. Acesso em: 30 maio 2017.

LEE, S.; KIM, B.; HUH, M.; Cho, S.; Park, S.; Lee, D. Mining transportation logs for understanding the after-assembly block manufacturing process in the shipbuilding industry. **Expert Systems with Applications**, v. 40, n. 1, p. 83–95, 2013. Elsevier Ltd. <<http://dx.doi.org/10.1016/j.eswa.2012.07.033>>.

LEHTO, T.; HINKKA, M.; HOLLMEN, J. Focusing Business Process Lead Time Improvements Using Influence Analysis. **CEUR Workshop Proceedings. Anais...**, 2017.

LENZ, R.; REICHERT, M. IT support for healthcare processes - premises, challenges, perspectives. **Data and Knowledge Engineering**, v. 61, n. 1, p. 39–58, 2007.

LEONI, M. DE; VAN DER AALST, W. M. P.; DEES, M. A general framework for correlating business process characteristics. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 8659 LNCS, p. 250–266, 2014. Disponível em: <[http://link.springer.com/chapter/10.1007/978-3-319-10172-9\\_16](http://link.springer.com/chapter/10.1007/978-3-319-10172-9_16)> <[http://link.springer.com/chapter/10.1007/978-3-319-10172-9\\_16](http://link.springer.com/chapter/10.1007/978-3-319-10172-9_16)> <<http://www.scopus.com/inward/record.url?eid=2-s2.0-84906729397%257B%257B%257D%257B%257D%257B%257D%257DpartnerID=40%257B%257B%257D%257B%257D%257B%257D%257D>>. Acesso em: 09 jan. 2018

MAITA, A. R. C. **Um estudo da aplicação de técnicas de inteligência computacional e de aprendizado em máquina de mineração de processos de negócio**, 2016. Departamento de Sistemas de Informação, Universidade de São Paulo. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/100/100131/tde-22012016-155157/publico/anarociodefesa.pdf>>. Acesso em: 19 nov. 2017

MAITA, A. R. C.; MARTINS, L. C.; LÓPEZ PAZ, C. R.; PERES, S. M.; FANTINATO, M. Process mining through artificial neural networks and support vector machines. **Business Process Management Journal**, v. 21, n. 6, p. 1391–1415, 2015. Disponível em: <<http://www.emeraldinsight.com/doi/10.1108/BPMJ-02-2015-0017>>. Acesso em: 19 nov. 2017

MANS, R. S.; SCHONENBERG, M. H.; SONG, M. S.; VAN DER AALST, W. M. P.; BAKKER, P. J. M. Application of Process Mining in Healthcare – A Case Study in a Dutch Hospital. **Proceedings of BIOSTEC 2008**, v. 25, p. 425–438, 2009. Disponível em: <[http://link.springer.com/chapter/10.1007/978-3-540-92219-3\\_32](http://link.springer.com/chapter/10.1007/978-3-540-92219-3_32)>. Acesso em: 31 jan. 2018

MARGHERITA, A. Business process management system and activities. **Business Process Management Journal**, v. 20, n. 5, p. 642–662, 2014. Disponível em: <<http://www.emeraldinsight.com/doi/10.1108/BPMJ-04-2013-0050>>. 28 out. 2017

MELDAU, D. C.; Hospitais Universitários. **InfoEscola**. Disponível em: <<https://www.infoescola.com/medicina/hospitais-universitarios/>>. Acesso em 20 dez. 2018

MENEGHINI, R.; PACKER, A. L. Is there science beyond English? **EMBO reports**, v. 8, n. 2, p. 112–116, 2007.

MEYER, A. D. Visual Data in Organizational Research. **Organization Science**, v. 2, n. 2, p. 218–236, 1991.



MIRELES, G.; ALBERTO, G.; JACOBO, J. R. Aplicacion del modelado de procesos en un curso de ingenieria de software. **Revista Electrónica de Investigación Educativa**, v. 3, n. 2, 2001.

MORAIS, R. M. DE; KAZAN, S.; PADUA, S. I. D. DE; COSTA, A. L. An analysis of BPM lifecycles: from a literature review to a framework proposal. **Business Process Management Journal**, v. 20, n. 3, 2014.

MÜNSTERMANN, B.; ECKHARDT, A.; WEITZEL, T. The performance impact of business process standardization. **Business Process Management Journal**, v. 16, n. 1, p. 29–56, 2010. Disponível em: <<http://www.emeraldinsight.com/doi/10.1108/14637151011017930>>. Acesso em: 13 nov. 2017

MYERS, D.; SURIADI, S.; RADKE, K.; FOO, E. Anomaly Detection for Industrial Control Systems using Process Mining. **Computers & Security**, 2018. Elsevier Ltd. <<https://doi.org/10.1016/j.cose.2018.06.002>>.

NOOIJEN, E. H. J.; DONGEN, B. F. VAN; FAHLAND, D. Automatic Discovery of Data-Centric and Artifact-Centric Processes. Business Process Management Workshops. **Anais...** p.316–327, 2013. Springer.

PEREIRA, G. B.; SANTOS, E. A. P.; MACENO, M. M. C. Aplicações e benefícios da mineração de processos na saúde. Anais do CBIS. **Anais...** 2018. Fortaleza: Sociedade Brasileira de Informática na Saúde.

PERIMAL-LEWIS, L.; QIN, S.; THOMPSON, C.; HAKENDORF, P. Gaining Insight from Patient Journey Data using a Process-Oriented Analysis Approach. In: K. Butler-Henderson; K. Gray (Orgs.); Health Informatics and Knowledge Management 2012. **Anais...** v. 129, p.59–66, 2012. Melbourne: Australian Computer Society Inc. Disponível em: <<http://crpit.com/confpapers/CRPITV129Perimal-Lewis.pdf>>. Acesso em: 31 maio 2017.

POELMANS, J.; DEDENE, G.; VERHEYDEN, G.; Van Der Musselle, H.; Viaene, S.; Peters, E. Combining business process and data discovery techniques for analyzing and improving integrated care pathways. In: P. Perner (Org.); Advances in Data Mining - Applications and Theoretical Aspects (ICDM 2010). **Anais...** p.505–517, 2010. Berlin: Springer. Disponível em: <<https://pdfs.semanticscholar.org/0d9d/ec2388ca7cad26d36e55c444db82784e6b12.pdf>>. Acesso em: 31 maio 2017.

PROPLAN, P. DE P. O. E F. Plano de desenvolvimento institucional 2017–2021. Disponível em: <[http://www.proplan.ufpr.br/portal/pdi/PDI UFPR 2017-2021.pdf](http://www.proplan.ufpr.br/portal/pdi/PDI%20UFPR%202017-2021.pdf)>. Acesso em: 8 dez. 2018.

RAMOS, L. T. **Healthcare process analysis: validation and improvements of a data-based method using process mining and visual analytics (MSc thesis)**, 2009. Eindhoven University of Technology. Disponível em: <<https://pure.tue.nl/ws/files/46955187/656484-1.pdf>>. Acesso em: 31 maio 2017.

REBUGE, Á.; FERREIRA, D. R. Business process analysis in healthcare environments: A methodology based on process mining. **Information Systems**, v. 37, n. 2, p. 99–116, 2012. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0306437911000044>>. Acesso em: 31 maio 2017.

RICHARDSON, R. J. **Pesquisa social: métodos e técnicas**. São Paulo: Atlas, 1999.

RIEMERS, P. **Process improvement in Healthcare: A data-based method using a combination of process mining and visual analytics**, 2009, Technology Management - Information Systems. Eindhoven University of Technology. Disponível em: <[http://alexandria.tue.nl/extra1/afstversl/tm/Riemers 2009.pdf](http://alexandria.tue.nl/extra1/afstversl/tm/Riemers%202009.pdf)>. Acesso em: 1 jun. 2017.

RIGGS, K. R.; REITMAN, Z. J.; MIELENZ, T. J.; GOODMAN, P. C. Relationship Between Time of First Publication and Subsequent Publication Success Among Non-PhD Physician-Scientists. **Journal of graduate medical education**, v. 4, n. 2, p. 196–201, 2012. Disponível em: <<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=23730441>>. Acesso em: 20 nov. 2017

RIVERA, F. A.; DURAN, A. Critical clouds and critical sets in resource-constrained projects. **International Journal of Project Management**, v. 22, p. 489–497, 2004.

RIZ, G.; SANTOS, E. A. P.; LOURES, E. D. F. R. Process mining to knowledge discovery in healthcare processes. **Advances in Transdisciplinary Engineering**, v. 4, p. 1019–1028, 2016.

ROJAS, E.; MUNOZ-GAMA, J.; SEPÚLVEDA, M.; CAPURRO, D. Process mining in healthcare: A literature review. **Journal of Biomedical Informatics**, v. 61, p. 224–236, 2016. Elsevier Inc. <<http://dx.doi.org/10.1016/j.jbi.2016.04.007>>. .

ROZINAT, A.; VAN DER AALST, W. M. P. Conformance checking of processes based on monitoring real behavior. **Information Systems**, v. 33, n. 1, p. 64–95, 2008.

SARAVANAN, M. S.; RAMA SREE, R. J. Analysis of Workflow and Process Mining in Dyeing Processing System. **International Journal of Computer Science and Technology**, v. 2, n. 3, p. 351–356, 2011.

SAYAL, M.; CASATI, F.; DAYAL, U.; SHAN, M.-C. Integrating workflow management systems with business-to-business interaction standards. **Proceedings 18th International Conference on Data Engineering**, p. 287–296, 2002. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=994737>>. Acesso em: 04 nov. 2017

SECRETARIA DE DESENVOLVIMENTO URBANO. Região Metropolitana de Curitiba. Disponível em:

<<http://www.comec.pr.gov.br/modules/conteudo/conteudo.php?conteudo=123>>. Acesso em: 8 dez. 2018.

SEETHAMRAJU, R.; MARJANOVIC, O. Role of process knowledge in business process improvement methodology: a case study. **Business Process Management Journal**, v. 15, n. 6, p. 920–936, 2009. Disponível em: <<http://www.emeraldinsight.com/doi/10.1108/14637150911003784>>. Acesso em: 31 out. 2017

SEIFY, M. Importance of KPI in BI system, Case study: Iranian industries. 2010 Seventh International Conference on Information Technology: New Generations. **Anais...** p.1245–1246, 2010. IEEE.

SHUKLA, N.; KEAST, J. E.; CEGLAREK, D. Improved workflow modelling using role activity diagram-based modelling with application to a radiology service case study. **Computer Methods and Programs in Biomedicine**, v. 116, n. 3, p. 274–298, 2014. Elsevier Ireland Ltd. <<http://dx.doi.org/10.1016/j.cmpb.2014.05.005>>.

SIDOROVA, A.; ISIK, O. Business process research: a cross-disciplinary review. **Business Process Management Journal**, v. 16, n. 4, p. 566–597, 2010. Disponível em: <<http://www.emeraldinsight.com/doi/10.1108/14637151011065928>>. Acesso em: 28 out. 2017

SILVA, E. L.; MENEZES, E. M. **Metodologia da Pesquisa e Elaboração de Dissertação**. 2005.

SMITH, A. **An Inquiry into the Nature and the Causes of the Wealth of Nations**. Oxford University Press, 1976.

SONG, M.; VAN DER AALST, W. M. P. Towards comprehensive support for organizational mining. **Decision Support Systems**, v. 46, n. 1, p. 300–317, 2008. Elsevier B.V. <<http://dx.doi.org/10.1016/j.dss.2008.07.002>>. .

STAAL, J. **Using Process and Data Improving Techniques to Define and Improve Standardization in a Healthcare Workflow Environment**, 2010, Eindhoven University of Technology. Disponível em: <[http://alexandria.tue.nl/extra2/afstversl/tm/Staal\\_2010.pdf](http://alexandria.tue.nl/extra2/afstversl/tm/Staal_2010.pdf)>. Acesso em: 07 jan. 2018

SURIADI, S.; WYNN, M. T.; OUYANG, C.; HOFSTEDE, A. H. M. TER; DIJK, N. VAN. Understanding process behaviours in a large insurance company in Australia: A case study. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). **Anais...** . v. 7908 LNCS, p.449–464, 2013. Valencia, Spain. Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-84879860686&partnerID=40&md5=61b72325da696888427087386f0ccf6a>>. Acesso em: 07 jan. 2018

TAYLOR, F. W. The principles of scientific management. **Management**, v. 6, p. 144, 1911. Disponível em: <<http://books.google.co.uk/books?id=tw-YGBovG8kC>>. 28 out. 2017

TRKMAN, P. The critical success factors of business process management. **International Journal of Information Management**, v. 30, n. 2, p. 125–134, 2010.

VIALE, P.; FRYDMAN, C.; PINATON, J. New Methodology for Modeling Large Scale Manufacturing Process: Using Process Mining Methods and Experts' Knowledge. 9th IEEE/ACS International Conference on Computer Systems and Applications (AICCSA). **Anais...** p.84–89, 2011.

WANG, F.; LEE, N.; HU, J.; Sun, J.; Ebadollahi, S.; Laine, A. F. A Framework for Mining Signatures from Event Sequences and its Applications in Healthcare Data. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 35, n. 2, p. 272, 2013. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6200289>>. Acesso em: 31 jan. 2018

WANG, J. Emergency healthcare workflow modeling and timeliness analysis. **IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans**, v. 42, n. 6, p. 1323–1331, 2012.

WEERDT, J. DE; SCHUPP, A.; VANDERLOOCK, A.; BAESENS, B. Process Mining for the multi-faceted analysis of business processes — A case study in a financial services organization. **Computers in Industry**, v. 64, n. 1, p. 57–67, 2013. Elsevier B.V. <<http://dx.doi.org/10.1016/j.compind.2012.09.010>>.

WENDLER, T.; LOEF, C. Workflow management-integration technology for efficient radiology. **Medicamundi**, n. November, p. 41–48, 2001. Disponível em: <[http://www.healthcare.philips.com/phpwc/main/about/assets/docs/medicamundi/mm\\_vol45\\_no4/mm\\_45-4\\_workflow.pdf](http://www.healthcare.philips.com/phpwc/main/about/assets/docs/medicamundi/mm_vol45_no4/mm_45-4_workflow.pdf)>. Acesso em: 04 nov. 2017

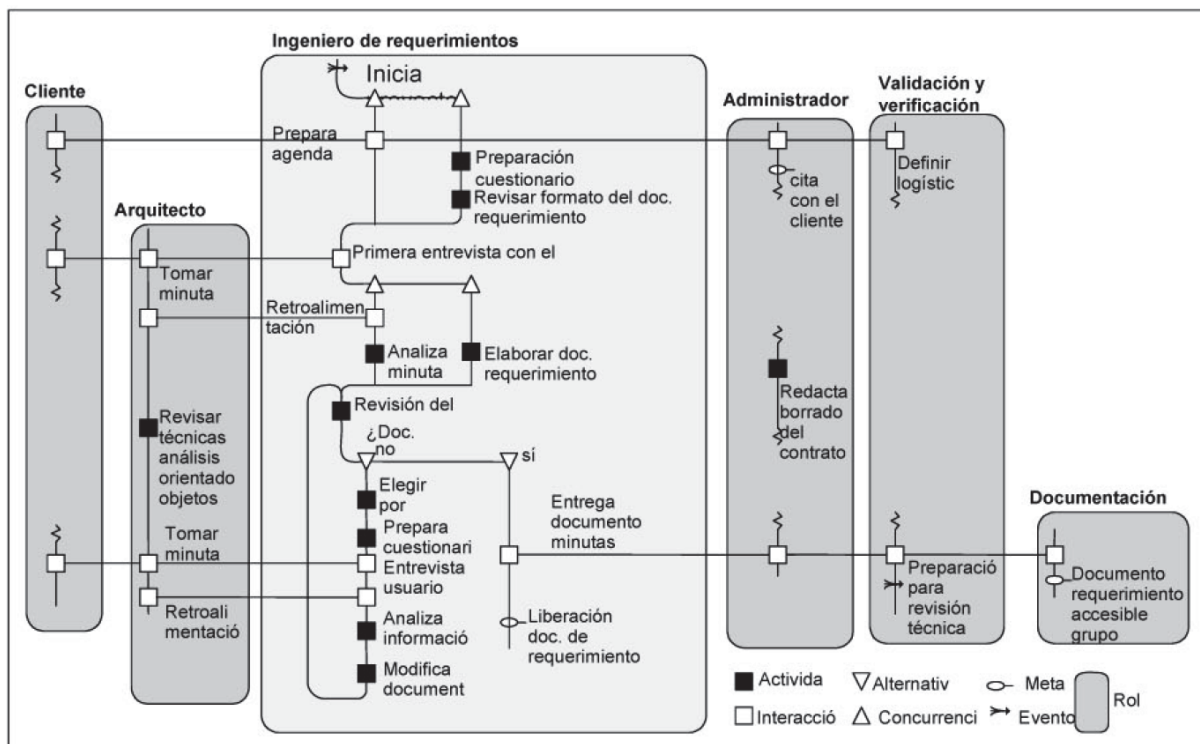
WIRTH, R.; HIPPEL, J. CRISP-DM: Towards a standard process model for data mining. **Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining**, n. 24959, p. 29–39, 2000.

XU, X.; JIN, T.; WEI, Z.; LV, C.; WANG, J. TCCPM: Topic-Based Clinical Pathway Mining. **Proceedings - 2016 IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2016**, p. 292–301, 2016.

XU, X.; JIN, T.; WEI, Z.; WANG, J. Incorporating Topic Assignment Constraint and Topic Correlation Limitation into Clinical Goal Discovering for Clinical Pathway Mining. **Journal of Healthcare Engineering**, v. 2017, 2017.

YANG, W.; SU, Q. Process Mining for Clinical Pathway: Literature Review and Future Directions. 11th International Conference on Service Systems and Service Management (ICSSSM). **Anais...** p.1–5, 2014. IEEE.

## ANEXO 1 - ROLE ACTIVITY DIAGRAM (RAD)



García Mireles, Gabriel Alberto, and Josefina Rodríguez Jacobo. "Aplicación del modelado de procesos en un curso de ingeniería de software." *Revista Electrónica de Investigación Educativa*, vol. 3, no. 2, 2001. *Academic OneFile*, <http://link-galegroup.ez22.periodicos.capes.gov.br/apps/doc/A146546490/AONE?u=capes&sid=AONE&xid=bc69f928>. Acceso em 23 Ago. 2018.